

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 15:05:44 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-1  
Perfect score: 50  
Sequence: 1 KKKKKKKKK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

- Database : SPTREMBL10.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

9/124280

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	730	3 Q99248	Q99248 saccharomyc
2	50	100.0	407	4 O00536	O00536 homo sapien
3	50	100.0	586	4 O95391	O95391 homo sapien
4	50	100.0	261	5 O18662	O18662 plasmodium
5	50	100.0	232	5 O76774	O76774 dictyosteli
6	50	100.0	580	5 O77334	O77334 plasmodium
7	50	100.0	4550	5 O77336	O77336 plasmodium
8	50	100.0	4981	5 O77372	O77372 plasmodium
9	50	100.0	2437	5 O77393	O77393 plasmodium
10	50	100.0	166	5 O77367	O77367 plasmodium
11	50	100.0	1139	5 O97242	O97242 plasmodium
12	50	100.0	644	5 O97258	O97258 plasmodium
13	50	100.0	872	5 O97296	O97296 plasmodium
14	50	100.0	756	10 O65716	O65716 arabidopsis
15	50	100.0	796	10 O49464	O49464 arabidopsis
16	50	100.0	215	11 O64075	O64075 rattus norv
17	50	100.0	129	11 O35807	O35807 rattus norv
18	47	94.0	483	5 O96148	O96148 plasmodium
19	47	94.0	972	5 O96233	O96233 plasmodium
20	47	94.0	2417	5 O97225	O97225 plasmodium
21	47	94.0	683	10 O64639	O64639 arabidopsis
22	47	94.0	530	11 O92202	O92202 mus musculu
23	46	92.0	404	1 O57810	O57810 pyrococcus
24	46	92.0	295	3 O12515	O12515 saccharomyc
25	46	92.0	254	5 O61095	O61095 trypanosoma
26	46	92.0	2010	5 O96171	O96171 plasmodium
27	46	92.0	1138	5 O96181	O96181 plasmodium
28	46	92.0	2013	5 O96216	O96216 plasmodium
29	46	92.0	3978	5 O97236	O97236 plasmodium

30	46	92.0	106	10 Q41092	Q41092 poncirus tr
31	46	92.0	202	10 Q41111	Q41111 phaseolus v
32	46	92.0	171	10 Q92R21	Q92R21 citrus unsh
33	46	92.0	473	11 O70396	O70396 mus musculu
34	46	92.0	496	11 O70310	O70310 mus musculu
35	46	92.0	1587	11 P70287	P70287 mus musculu
36	46	92.0	1560	11 P97868	P97868 mus musculu
37	46	92.0	376	11 O61048	O61048 mus musculu
38	46	92.0	1630	13 Q90753	Q90753 gallus gall
39	45	90.0	1051	3 Q01694	Q01694 pneumocysti
40	45	90.0	790	3 Q02744	Q02744 pneumocysti
41	45	90.0	72	3 P89493	P89493 saccharomyc
42	45	90.0	886	4 Q15361	Q15361 homo sapien
43	45	90.0	328	5 O15798	O15798 p strain dd
44	45	90.0	368	5 O45198	O45198 caenorhabdi
45	45	90.0	861	5 Q21571	Q21571 caenorhabdi

ALIGNMENTS

RESULT 1  
Q99248  
ID Q99248 PRELIMINARY; PRT: 730 AA.  
AC Q99248;  
DT 01-NOV-1996 (TREMREL. 01; Created)  
DT 01-NOV-1996 (TREMREL. 01; Last sequence update)  
DT 01-NOV-1998 (TREMREL. 08; Last annotation update)  
DE CHROMOSOME XY READING FRAME ORF YOR019W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DE HAAN M., GRIVELL L.A., MAARSE A.C.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA STRAIN-FY1679;  
RC DE HAAN M., MAARSE A.C., GRIVELL L.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN-FY1679;  
RX MEDLINE: 94019318.  
RA DUMONT M.E., SCHLICHTER J.B., CARDILLO T.S., HAYES M.K., BETHLENDY G.,  
RA SHERMAN F.;  
RT "CYC2 encodes a factor involved in mitochondrial import of yeast  
RT cytochrome c";  
RL Mol. Cell. Biol. 13:6442-6451(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RX MEDLINE: 94169519.  
RA LEE Y.S., SHIMIZU J., YODA K., YAMASAKI M.;  
RT "Molecular cloning of a gene, DHS1, which complements a  
RT drug-hypersensitive mutation of the yeast Saccharomyces cerevisiae";  
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).  
DR EMBL; X74927; CAA99209.1; -  
DR EMBL; X87331; CAA60768.1; -  
SQ SEQUENCE 730 AA; 83365 MW; 09D72CD9 CRC32;

Query Match 100.0%; Score 50; DB 3; Length 730;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKKKKKKKK 10  
|||||

us-09-124-280a-1.rspt

Wed Sep 8 09:41:58 1999

Db 713 KKKKKKKK 722

RESULT 2

ID 000536 PRELIMINARY; PRT; 407 AA.

AC 000536;

DT 01-JUL-1997 (TREMELrel. 04, Created)

DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)

DE TTF-1 INTERACTING PEPTIDE 5 (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.

EN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;

RA JANSAP, GRUMMT I.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF000422; AAB60864.1;

DR PFAM; PF01429; MBD; 1.

FT NON TER 1 407

FT NON TER 407 407

SQ SEQUENCE 407 AA; 44759 MW; 57F34A7F CRC32;

Query Match 100.0%; Score 50; DB 4; Length 407;

Best Local Similarity 100.0%; Pred. No. 4.2;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10

DB 398 KKKKKKKK 407

RESULT 3

ID 095391 PRELIMINARY; PRT; 586 AA.

AC 095391;

DT 01-MAY-1999 (TREMELrel. 10, Created)

DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)

DE STEP II SPLICING FACTOR SLU7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.

EN [1]

RP SEQUENCE FROM N.A.

RA CHUA K., REED R.;

RT "A human step II splicing factor hslu7 involved in restructuring the spliceosome between the catalytic steps of the pre-mRNA splicing reaction."

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF101074; AAD13774.1;

SQ SEQUENCE 586 AA; 66356 MW; E67A16B7 CRC32;

Query Match 100.0%; Score 50; DB 4; Length 586;

Best Local Similarity 100.0%; Pred. No. 5.4;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10

DB 498 KKKKKKKK 507

RESULT 4

ID 018662 PRELIMINARY; PRT; 261 AA.

AC 018662;

DT 01-JAN-1998 (TREMELrel. 05, Created)

DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)

DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)

CG9 PROTEIN.

GN CG9.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DD2, AND HB3;

RX MEDLINE: 98054002.

RA SU X.-Z., KIRKMAN L.A., FUJIOKA H., WELLEMS T.E.;

RT "Complex polymorphisms in an approximately kDa protein are linked to chloroquine-resistant P. falciparum in Southeast Asia and Africa."

RL Cell 91:593-603(1997).

EN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=DD2, AND HB3;

RA SU X.-Z., KIRKMAN L.A., WELLEMS T.E.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF030694; AAC47839.1;

DR EMBL; AF030690; AAC47852.1;

SQ SEQUENCE 261 AA; 31108 MW; 5CFBDC37 CRC32;

Query Match 100.0%; Score 50; DB 5; Length 261;

Best Local Similarity 100.0%; Pred. No. 3.1;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10

DB 231 KKKKKKKK 240

RESULT 5

ID 076774 PRELIMINARY; PRT; 232 AA.

AC 076774;

DT 01-NOV-1998 (TREMELrel. 08, Created)

DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)

DE GTP BINDING PROTEIN RARE7L.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Dictyostellida; Dictyostelium.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX2;

RA CLOUGHERTY C., GERISCH G.;

RT "Characterization of a Rab-like GTPase from Dictyostelium discoideum."

RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF081681; AAC34837.1;

DR PFAM; PF00071; ras; 1.

SQ SEQUENCE 232 AA; 27123 MW; 472785B4 CRC32;

Query Match 100.0%; Score 50; DB 5; Length 232;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10

DB 207 KKKKKKKK 216

RESULT 6

ID 077334 PRELIMINARY; PRT; 580 AA.

AC 077334;

DT 01-NOV-1998 (TREMELrel. 08, Created)

DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)

DE MAL3P3.11 PROTEIN.

GN MAL3P3.11.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.



```

RN  [1]
RP  SEQUENCE FROM N.A.
RA  MUNGALL K., LAWSON D., BARRELL B.;
RL  Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL: Z98547; CAB1119.1; -.
DR  PFAM: PF00782; DSPC.1.
SQ  SEQUENCE 580 AA; 68768 MW;  B5486486 CRC32;

Query Match      100.0%; Score 50; DB 5; Length 580;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KKKKKKKKK 10
    |
DB  543 KKKKKKKKK 552

RESULT 7
O77336 PRELIMINARY; PRT; 4550 AA.
AC O77336;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MAL3P3.18 PROTEIN.
GN MAL3P3.18.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA MUNGALL K., LAWSON D., BARRELL B.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98547; CAB11121.1; -.
SQ SEQUENCE 4550 AA; 534109 MW;  CC69B096 CRC32;

Query Match      100.0%; Score 50; DB 5; Length 4550;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KKKKKKKKK 10
    |
DB  162 KKKKKKKKK 171

RESULT 8
O77372 PRELIMINARY; PRT; 4981 AA.
AC O77372;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MAL3P6.23 PROTEIN.
GN MAL3P6.23.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA MURPHY L., LAWSON D., BARRELL B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98551; CAB11128.1; -.
SQ SEQUENCE 4981 AA; 593251 MW;  183E8965 CRC32;

Query Match      100.0%; Score 50; DB 5; Length 4981;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KKKKKKKKK 10
    |
DB  1434 KKKKKKKKK 1443
```

```

RESULT 9
O77393 PRELIMINARY; PRT; 2437 AA.
AC O77393;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MAL3P6.2 PROTEIN.
GN MAL3P6.2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA MURPHY L., LAWSON D., BARRELL B.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98551; CAB11149.1; -.
SQ SEQUENCE 2437 AA; 295348 MW;  E6280382 CRC32;

Query Match      100.0%; Score 50; DB 5; Length 2437;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KKKKKKKKK 10
    |
DB  1386 KKKKKKKKK 1395

RESULT 10
O77367 PRELIMINARY; PRT; 166 AA.
AC O77367;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE MAL3P6.28 PROTEIN.
GN MAL3P6.28.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA MURPHY L., LAWSON D., BARRELL B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98551; CAB11123.2; -.
SQ SEQUENCE 166 AA; 19788 MW;  8FF3D037 CRC32;

Query Match      100.0%; Score 50; DB 5; Length 166;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KKKKKKKKK 10
    |
DB  49 KKKKKKKKK 58

RESULT 11
O97242 PRELIMINARY; PRT; 1139 AA.
AC O97242;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE MAL3P2.21 PROTEIN.
GN MAL3P2.21.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-3D7;  
RA CHURCHER C., BOWMAN S., LAWSON D., QUAIL M., BARRELL B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL034558; CAB39008.1; -;  
SQ SEQUENCE 1139 AA; 135437 MW; C2F82EA2 CRC32;  
  
Query Match 100.0%; Score 50; DB 5; Length 1139;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKKKKKKKK 10  
DB 875 KKKKKKKKK 884  
|||||  
  
RESULT 12  
O97258 PRELIMINARY; PRT; 644 AA.  
AC O97258;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE MAL3P5.1 PROTEIN.  
GN MAL3P5.1.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-3D7;  
RA BOWMAN S., SKELTON J., CHURCHER C., LAWSON D., QUAIL M., BARRELL B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL034556; CAB38969.1; -;  
SQ SEQUENCE 644 AA; 78302 MW; 04F50B4F CRC32;  
  
Query Match 100.0%; Score 50; DB 5; Length 644;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKKKKKKKK 10  
DB 100 KKKKKKKKK 109  
|||||  
  
RESULT 13  
O97296 PRELIMINARY; PRT; 872 AA.  
AC O97296;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE MAL3P7.30 PROTEIN.  
GN MAL3P7.30.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-3D7;  
RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL034559; CAB39043.1; -;  
SQ SEQUENCE 872 AA; 105284 MW; 33B75496 CRC32;

Query Match 100.0%; Score 50; DB 5; Length 872;  
Best Local Similarity 100.0%; Pred. No. 7.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKKKKKKKK 10  
DB 420 KKKKKKKKK 429  
|||||

RESULT 14  
O65716 PRELIMINARY; PRT; 756 AA.  
AC O65716;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE HYPOTHETICAL 85.8 KD PROTEIN.  
GN T5K18.240.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BEVAN M., VAN DER SCHUEREN J., CHUANG Y.-J., VOET M., ROBBERN J.,  
RA VOLCKAERT G., BANCROFT I., MEWES H.W., MAYER K.F.X., SCHUELLER C.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU ARABIDOPSIS SEQUENCING PROJECT;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL022580; CAA18633.1; -;  
DR PFAM; PF00534; Glycos\_transf\_1; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 756 AA; 85770 MW; 6C2625D4 CRC32;  
  
Query Match 100.0%; Score 50; DB 10; Length 756;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKKKKKKKK 10  
DB 739 KKKKKKKKK 748  
|||||  
  
RESULT 15  
O49464 PRELIMINARY; PRT; 796 AA.  
AC O49464;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE HYPOTHETICAL 90.5 KD PROTEIN.  
GN F24J7.20.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BEVAN M., VITALE D., LIGUORI R., ARGIRIOU A., DE SIMONE V.,  
RA BANCROFT I., MEWES H.W., MAYER K., SCHUELLER C.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL021768; CAA16923.2; -;  
DR PFAM; PF00534; Glycos\_transf\_1; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 796 AA; 90512 MW; F3B89BF1 CRC32;  
  
Query Match 100.0%; Score 50; DB 10; Length 796;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KKKKKKKKK 10  
DB 779 KKKKKKKKK 788  
|||||

Wed Sep 8 09:41:58 1999

us-09-124-280a-1.rspt

Page 5

Search completed: September 7, 1999, 20:34:23  
Job time: 19719 sec

**THIS PAGE BLANK (USPTO)**

AS MEDLINE, 50105552.

RA MALLET L., BUSSEAU F., JACQUET M.;  
 RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
 RT MEF2, CAP/SRV2, NAM9, FKBI/FPRI/RBP1, MOM22 and CPT1, predicts an  
 RT adenosine deaminase gene and 14 new open reading frames.";   
 RL YEAST 11:1195-1209(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; 246843; G854495; -  
 DR EMBL; 271419; E239951; -  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 130 AA; 15319 MW; 45AD0B99 CRC32;

Query Match 100.0%; Score 50; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 0.45;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 Db 49 KKKKKKKKK 58

RESULT 3  
 AMP2\_HUMAN  
 ID AMP2\_HUMAN STANDARD; PRT; 478 AA.  
 AC P50579;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METHIONINE AMINOPEPTIDASE 2 (EC 3.4.11.18) (METAP 2) (PEPTIDASE M 2)  
 DE (INITIATION FACTOR 2 ASSOCIATED 67 KD GLYCOPROTEIN) (P67).  
 GN P67EIF2.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 95372350.  
 RA APIN S.M., KENDALL R.L., HALL L., WEAVER L.H., STEWART A.E.,  
 RA MATTHEWS B.W., BRADSHAW R.A.;  
 RT "Eukaryotic methionyl aminopeptidases: two classes of  
 RL cobalt-dependent enzymes.";   
 RL PROC. NATL. ACAD. SCI. U.S.A. 92:7714-7718(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LIVER;  
 RX MEDLINE; 95178556.  
 RA LI X., CHANG Y.;  
 RA "Molecular cloning of a human complementary DNA encoding an  
 RT initiation factor 2-associated protein (p67).";  
 RL BIOCHIM. BIOPHYS. ACTA 1260:333-336(1995).  
 CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT  
 CC PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: L-METHIONYLPEPTIDE + H(2)O = L-METHIONINE  
 CC + PEPTIDE.  
 CC -1- COFACTOR: COBALT.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24C; ALSO KNOWN AS THE  
 CC MAP FAMILY 2.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U29607; G903982; -  
 DR EMBL; U13261; G687243; -  
 DR MIM; 601870; -  
 DR PROSITE; PS01202; MAP\_2; 1.  
 DR PFAM; PF00557; pep\_M24; 1.  
 DR HSP; P56218; IXGO.  
 KW HYDROLASE; AMINOPEPTIDASE; COBALT.  
 FT DOMAIN 36 45 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 82 93 ASP/GLU-RICH (ACIDIC).  
 FT METAL 98 106 POLY-LYS.  
 FT METAL 251 251 COBALT (BY SIMILARITY).  
 FT METAL 262 262 COBALT (BY SIMILARITY).  
 FT METAL 331 331 COBALT (BY SIMILARITY).  
 FT METAL 364 364 COBALT (BY SIMILARITY).  
 FT METAL 460 460 COBALT (BY SIMILARITY).  
 SQ SEQUENCE 478 AA; 52891 MW; 18AFAE0A CRC32;

Query Match 94.0%; Score 47; DB 1; Length 478;  
 Best Local Similarity 90.0%; Pred. No. 3.3;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 Db 98 KKKKKKKKK 107

RESULT 4  
 AMP2\_MOUSE  
 ID AMP2\_MOUSE STANDARD; PRT; 478 AA.  
 AC Q08663;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE METHIONINE AMINOPEPTIDASE 2 (EC 3.4.11.18) (METAP 2) (PEPTIDASE M 2)  
 DE (INITIATION FACTOR 2 ASSOCIATED 67 KD GLYCOPROTEIN) (P67).  
 GN P67EIF2.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=LIVER;  
 RA SEKIGUCHI S., SUZUKI E.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT  
 CC PROTEINS (BY SIMILARITY).  
 CC -1- FUNCTION: PROTECTS EIF-2 ALPHA-SUBUNIT FROM INHIBITORY  
 CC PHOSPHORYLATION BY EIF-2 KINASES. PLAYS A CRITICAL ROLE IN THE  
 CC REGULATION OF PROTEIN SYNTHESIS. IT ALSO INTERACTS WITH THE  
 CC EIF-2 GAMMA-SUBUNIT (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: L-METHIONYLPEPTIDE + H(2)O = L-METHIONINE  
 CC + PEPTIDE.  
 CC -1- COFACTOR: COBALT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24C; ALSO KNOWN AS THE  
 CC MAP FAMILY 2.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB003144; D1020571; -  
 DR PROSITE; PS01202; MAP\_2; 1.  
 DR PFAM; PF00557; pep\_M24; 1.  
 DR HSP; P56218; IXGO.  
 KW HYDROLASE; AMINOPEPTIDASE; COBALT.  
 FT DOMAIN 36 46 ARG/LYS-RICH (BASIC).  
 FT

FT DOMAIN 82 93 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 98 106 POLY-LYS.  
 FT METAL 251 251 COBALT (BY SIMILARITY).  
 FT METAL 262 262 COBALT (BY SIMILARITY).  
 FT METAL 331 331 COBALT (BY SIMILARITY).  
 FT METAL 364 364 COBALT (BY SIMILARITY).  
 FT METAL 460 460 COBALT (BY SIMILARITY).  
 SQ SEQUENCE 478 AA; 52921 MW; 119FCE5 CRC32;

Query Match 94.0%; Score 47; DB 1; Length 478;  
 Best Local Similarity 90.0%; Pred. No. 3.3;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10  
 Db 98 KKKKKKKKKR 107

## RESULT 5

AMP2\_RAT STANDARD; PRT; 478 AA.  
 AC P38062;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DE METHIONINE AMINOPEPTIDASE 2 (EC 3.4.11.18) (METAP 2) (PEPTIDASE M 2)  
 DE (INITIATION FACTOR 2 ASSOCIATED 67 KD GLYCOPROTEIN) (P67).  
 GN P67EIF2.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-REUBER H35; TISSUE=LIVER;  
 RX MEDLINE; 93266517.  
 RA WU S., GUPTA S., CHATTERJEE N., HILEMAN R.E., KINZY T.G.,  
 RA DENLOW N.D., MERRICK W.C., CHAKRABARTI D., OSTERMAN J.C., GUPTA N.K.;  
 RT "Cloning and characterization of complementary DNA encoding the  
 RT eukaryotic initiation factor 2-associated 67-kDa protein (p67).";  
 RL J. BIOL. CHEM. 268:10796-10801(1993).  
 RN [2]  
 RP REVISIONS TO C-TERMINAL.  
 RX MEDLINE; 95372350.  
 RA ARFIN S.M., KENDALL R.L., HALL L., WEAVER L.H., STEWART A.E.,  
 RA MATTHEWS B.W., BRADSHAW R.A.;  
 RT "Eukaryotic methionyl aminopeptidases: two classes of  
 RT cobalt-dependent enzymes.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 92:7714-7718(1995).  
 CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT  
 CC PROTEINS.  
 CC -1- FUNCTION: PROTECTS EIF-2 ALPHA-SUBUNIT FROM INHIBITORY  
 CC PHOSPHORYLATION BY EIF-2 KINASES. PLAYS A CRITICAL ROLE IN THE  
 CC REGULATION OF PROTEIN SYNTHESIS. IT ALSO INTERACTS WITH THE  
 CC EIF-2 GAMMA-SUBUNIT.  
 CC -1- CATALYTIC ACTIVITY: L-METHIONYLPEPTIDE + H(2)O -> L-METHIONINE  
 CC + PEPTIDE.  
 CC -1- COFACTOR: COBALT (BY SIMILARITY).  
 CC -1- PTM: CONTAINS 12 O-LINKED GLCNAC.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24C; ALSO KNOWN AS THE  
 CC MAP FAMILY 2.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; L10652; G204004; -.  
 DR PIR; A46702; A46702.  
 DR PROSITE; PS01202; MAP\_2; 1.

DR PRAM; PF00557; pep\_M24; 1.  
 DR HSSP; P56218; LXGO.

KW HYDROLASE; AMINOPEPTIDASE; COBALT; GLYCOPROTEIN.  
 FT DOMAIN 36 46 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 82 93 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 98 106 POLY-LYS.  
 FT METAL 251 251 COBALT (BY SIMILARITY).  
 FT METAL 262 262 COBALT (BY SIMILARITY).  
 FT METAL 331 331 COBALT (BY SIMILARITY).  
 FT METAL 364 364 COBALT (BY SIMILARITY).  
 FT METAL 460 460 COBALT (BY SIMILARITY).  
 FT CONFLICT 464 478 LRPTCKEVVSRGDDY -> CAQPVKKLSAEEMTIKT (IN  
 REF. 1).  
 SQ SEQUENCE 478 AA; 53052 MW; 1AA3A160 CRC32;

Query Match 94.0%; Score 47; DB 1; Length 478;  
 Best Local Similarity 90.0%; Pred. No. 3.3;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10  
 Db 98 KKKKKKKKKR 107

## RESULT 6

YHP5\_YEAST STANDARD; PRT; 144 AA.  
 AC P38808;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE HYPOTHEICAL 16.7 KD PROTEIN IN HXT1-HXT5 INTERGENIC REGION.  
 GN YHR095W.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RX MEDLINE; 94378003.  
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,  
 RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,  
 RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,  
 RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,  
 RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,  
 RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,  
 RA VAUDIN M.;  
 RT \*Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII.\*;  
 RL SCIENCE 265:2077-2082(1994).

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL; U00060; G487930; -.  
 DR PIR; S46716; S46716.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 144 AA; 16678 MW; 91D97E67 CRC32;

Query Match 94.0%; Score 47; DB 1; Length 144;  
 Best Local Similarity 90.0%; Pred. No. 1.2;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10  
 Db 111 KKKKKKKKKR 120

```

RESULT 7
ID PHPA_PLACH STANDARD; PRT; 441 AA.
AC Q02752;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE ACIDIC PHOSPHOPROTEIN PRECURSOR (50 KD ANTIGEN).
GN PCWAL
OS PLASMIDIUM CHABAUDI.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IP-PC1;
RX MEDLINE: 93116806.
RA DELERSNIJDER W., PRASOMSIITI P., TUNGPRADURKUL S., HENDRIX D.,
RA HAMERS-CASTERMAN C., HAMERS R.;
RT "Structure of a Plasmidium chabaudi acidic phosphoprotein that is
RT associated with the host erythrocyte membrane.";
RL MOL. BIOCHEM. PARASITOL. 56:59-68(1992).
CC -!- FUNCTION: DURING INFECTION, THIS PHOSPHOPROTEIN PROBABLY MODULATES
CC THE STRUCTURE OF THE RED CELL MEMBRANE TO THE ADVANTAGE OF THE
CC PARASITE, ALTHOUGH ITS PRECISE FUNCTION IS NOT KNOWN.
CC -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN ON THE
CC CYTOPLASMIC FACE OF THE HOST ERYTHROCYTE MEMBRANE.
CC -!- ASSOCIATED WITH THE HOST RED CELL MEMBRANE THROUGHOUT THE ENTIRE
CC ERYTHROCYTIC CYCLE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M95789; G160603; -.
DR PIR: A48455; A48455.
KW PHOSPHORYLATION; SIGNAL; ANTIGEN; MEMBRANE; REPEAT; ERYTHROCYTE.
FT SIGNAL 1 15
FT CHAIN 16 441
FT DOMAIN 186 313
FT REPEAT 186 193 1-1.
FT REPEAT 194 201 1-2.
FT REPEAT 202 209 1-3.
FT REPEAT 210 217 1-4.
FT REPEAT 218 225 1-5.
FT REPEAT 226 233 1-6.
FT REPEAT 234 241 1-7.
FT REPEAT 242 249 1-8.
FT REPEAT 250 257 1-9.
FT REPEAT 258 265 1-10.
FT REPEAT 266 273 1-11.
FT REPEAT 274 281 1-12.
FT REPEAT 282 289 1-13.
FT REPEAT 290 297 1-14.
FT REPEAT 298 305 1-15.
FT REPEAT 306 313 1-16.
FT DOMAIN 353 370 2 X 9 AA TANDEM REPEATS.
FT REPEAT 353 360 2-1.
FT REPEAT 361 368 2-2.
FT REPEAT 371 417 LYS-RICH (BASIC).
FT CARBOHYD 21 21 POTENTIAL.
FT CARBOHYD 112 112 POTENTIAL.
SQ SEQUENCE 441 AA; 49708 MW; 757C1E17 CRC32;

```

Query Match 92.0%; Score 46; DB 1; Length 441;  
 Best Local Similarity 90.0%; Pred. No. 4.1;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KKKKKKKK 10
Db 394 KKKKKKKK 403

RESULT 8
ID SN24_HUMAN STANDARD; PRT; 1647 AA.
AC P51532;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE POSSIBLE GLOBAL TRANSCRIPTION ACTIVATOR SNF2L4 (SNF2-BETA) (BRG-1
DE PROTEIN) (MITOTIC GROWTH AND TRANSCRIPTION ACTIVATOR) (BRAHMA
DE PROTEIN HOMOLOGY 1).
GN SNF2L4 OR BRG1 OR SNF2B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94030144.
RA KHAVARI P.A., PETERSON C.L., TAMKUN J.W., MENDEL D.B., CRABTREE G.R.;
RT "BRG1 contains a conserved domain of the SWI2/SNF2 family necessary
RT for normal mitotic growth and transcription.";
RL NATURE 366:170-174(1993).
RN [2]
RP REVISIONS.
RA KHAVARI P.A., PETERSON C.L., TAMKUN J.W., MENDEL D.B., CRABTREE G.R.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDJB DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94268902.
RA CHIBA H., MURAMATSU M., NOMOTO A., KATO H.;
RT "Two human homologues of Saccharomyces cerevisiae SWI2/SNF2 and
RT Drosophila brahma are transcriptional coactivators cooperating with
RT the estrogen receptor and the retinoic acid receptor.";
RL NUCLEIC ACIDS RES. 22:1815-1820(1994).
CC -!- FUNCTION: TRANSCRIPTIONAL COACTIVATOR COOPERATING WITH NUCLEAR
CC HORMONE RECEPTORS TO POTENTIATE TRANSCRIPTIONAL ACTIVATION.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U29175; G902046; -.
DR EMBL: D26156; G505088; -.
DR PROSITE: PS00633; BROMODOMAIN_1; 1.
DR PROSITE: PS00014; BROMODOMAIN_2; 1.
DR PFAM: PF00176; SNF2_N; 1.
DR PFAM: PF00271; helicase_C; 1.
DR PFAM: PF00439; bromodomain; 1.
KW TRANSCRIPTION REGULATION; NUCLEAR PROTEIN; ACTIVATOR; BROMODOMAIN;
ATP-BINDING; HELICASE.
FT DOMAIN 578 588 POLY-LYS.
FT DOMAIN 663 672 POLY-GLU.
FT NP_BIND 779 786 ATP (POTENTIAL).
FT SITE 881 884 DEGH BOX.
FT DOMAIN 1360 1364 POLY-GLU.
FT DOMAIN 1471 1547 BROMODOMAIN.
FT DOMAIN 1571 1584 POLY-GLU.
SQ SEQUENCE 1647 AA; 184585 MW; 6D22A23A CRC32;

```

Query Match 92.0%; Score 46; DB 1; Length 1647;  
 Best Local Similarity 90.0%; Pred. No. 13;



Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
I:|||||

DB 578 KKKKKKKKK 587

# RESULT 9

ID T2D2\_DROME STANDARD; PRT: 1213 AA.

AC Q24325;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 13-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE TRANSCRIPTION INITIATION FACTOR TFIID 150 KD SUBUNIT (TAFII-150)

DE (TAFII150).

GN TAFI150.

OS DROSOPHILA MELANOGASTER (FRUIT FLY).

OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;

OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;

OC DROSOPHILIDAE; DROSOPHILA.

[1]

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP TISSUE-EMBRYO.

RC MEDLINE; 94233377.

RX VERRIJZER C.P., YOKOMORI K., CHEN J.-L., TJIAN R.;

RT "Drosophila TAFII150: similarity to yeast gene TSM-1 and specific

binding to core promoter DNA.";

RL SCIENCE 264:933-941(1994).

CC -1- FUNCTION: TAFS ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID

(TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA

POLYMERASE TRANSCRIPTION. TAFII-150 IS AN ESSENTIAL SUBUNIT WHICH

INTERACTS DIRECTLY WITH TBP AND TAFII-250 AND BINDS TO CORE

PROMOTOR DNA.

CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A

NUMBER OF TBP-ASSOCIATED FACTORS (TAFS).

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- SIMILARITY: TO YEAST TAFII-150 (TSM1).

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: X79243; G541665; -.

DR FLYBASE; FBgn0011836; Tafi150.

DR TRANSFAC; T02120; -.

KW TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.

FT DOMAIN 845 1213 BINDS TO TBP AND TAFII-250.

FT DOMAIN 1138 1183 HIGHLY CHARGED.

SQ SEQUENCE 1213 AA; 138533 MW; 2A07DC14 CRC32;

Query Match 92.0%; Score 46; DB 1; Length 1213;

Best Local Similarity 90.0%; Pred. No. 9.9;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
I:|||||

DB 1150 KKKKKKKKK 1159

# RESULT 10

ID YK20\_YEAST STANDARD; PRT: 167 AA.

AC P36133;

DT 01-JUN-1994 (REL. 29, CREATED)

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

DT 13-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 19.2 KD PROTEIN IN GAP1-NAP1 INTERGENIC REGION.

GN YKR040C.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

OC SACCHAROMYCETACEAE; SACCHAROMYCES.

[1]

RP SEQUENCE FROM N.A.

RA URRESTARAZU L.A., JAUNIAUX J.-C.;

RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL: Z28255; G486481; -.

DR EMBL; Z28255; G486481; -.

DR PIR; S38112; S38112.

KW HYPOTHETICAL PROTEIN.

FT DOMAIN 144 167 LYS-RICH (HIGHLY BASIC).

SQ SEQUENCE 167 AA; 19222 MW; 9AC0E484 CRC32;

Query Match 92.0%; Score 46; DB 1; Length 167;

Best Local Similarity 90.0%; Pred. No. 1.8;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
I:|||||

DB 143 EKKKKKKKK 152

# RESULT 11

ID AMP2\_YEAST STANDARD; PRT: 421 AA.

AC P38174;

DT 01-OCT-1994 (REL. 30, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE METHIONINE AMINOPEPTIDASE 2 (EC 3.4.11.18) (METAP 2) (PEPTIDASE M 2).

GN MAP2 OR YBL091C OR YBL0701.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

OC SACCHAROMYCETACEAE; SACCHAROMYCES.

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=YPH500;

RX MEDLINE; 96109265.

RA LI X., CHANG Y.-H.;

RT "Amino-terminal protein processing in *Saccharomyces cerevisiae* is an

essential function that requires two distinct methionine

aminopeptidases.";

RT PROC. NATL. ACAD. SCI. U.S.A. '92:12357-12361(1995).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RX MEDLINE; 96076635.

RA OBERMAIER B., GASSENHUBER J., PIRAVANDI E., DOMDEY H.;

RT "Sequence analysis of a 78.6 kb segment of the left end of

*Saccharomyces cerevisiae* chromosome II.";

RT YEAST 11:1103-1112(1995).

CC -1- FUNCTION: REMOVES THE AMINO-TERMINAL METHIONINE FROM NASCENT

PROTEINS.

CC -1- CATALYTIC ACTIVITY: L-METHIONYLPEPTIDE + H(2)O = L-METHIONINE

+ PEPTIDE.

CC -1- COFACTOR: COBALT (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24C; ALSO KNOWN AS THE

NAP FAMILY 2.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: U17437; G1045302; -  
 DR EMBL: X79489; G496684; ALT\_INIT.  
 DR EMBL: Z35852; E304508; -  
 DR PIR: S45411; S45411.  
 DR SGD: L0003044; MAP2.  
 DR PROSITE: PS01202; MAP\_2; 1.  
 DR PFAM: PF00557; pep\_M24; 1.  
 DR HSSP: P56218; 1XGO.  
 KW HYDROLASE; AMINOPEPTIDASE; COBALT.  
 FT METAL 194 194 COBALT (BY SIMILARITY).  
 FT METAL 205 205 COBALT (BY SIMILARITY).  
 FT METAL 274 274 COBALT (BY SIMILARITY).  
 FT METAL 307 307 COBALT (BY SIMILARITY).  
 FT METAL 403 403 COBALT (BY SIMILARITY).  
 FT CONFLICT 86 86 D -> V (IN REF. 2).  
 SQ SEQUENCE 421 AA; 47518 MW; 361C0A4E CRC32;

Query Match 90.0%; Score 45; DB 1; Length 421;  
 Best Local Similarity 90.0%; Pred. No. 5.3;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10  
 Db 41 KKKKKKKKK 50

## RESULT 12

EF1H\_XENLA STANDARD; PRT; 437 AA.  
 AC Q91375;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE ELONGATION FACTOR 1-GAMMA TYPE 2 (EF-1-GAMMA) (P47).  
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 CC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OOCYTE;  
 RA MORALES J., BASSEZ T., CORMIER P., MULNER-LORILLON O., BELLE R.,  
 RA OSBORNE H.B.;  
 RT "Expression of elongation factor 1 alpha (EF-1 alpha) and 1 beta  
 gamma (EF-1 beta gamma) are uncoupled in early xenopus embryos.";  
 RL DEV. GENET. 14:440-448(1993).  
 CC -!- FUNCTION: PROBABLY PLAYS A ROLE IN ANCHORING THE COMPLEX TO OTHER  
 CC CELLULAR COMPONENTS.  
 CC -!- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
 CC DELTA, AND GAMMA.  
 CC -!- SIMILARITY: TO OTHER SPECIES' GAMMA SUBUNIT AS WELL AS TO THE  
 CC N-TERMINAL OF SOME EUKARYOTIC AMINO ACID-TRNA SYNTHETASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: S69726; G545507; -  
 DR PROSITE: PS50040; EF1G; 1.  
 DR PFAM: PF00043; gluts; 1.  
 DR PFAM: PF00647; EF1G\_domain; 1.  
 KW ELONGATION FACTOR; PROTEIN BIOSYNTHESIS.

SQ SEQUENCE 437 AA; 50248 MW; 3D632B6E CRC32;

Query Match 90.0%; Score 45; DB 1; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 5.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 9  
 Db 242 KKKKKKKKK 250

## RESULT 13

YMH8\_YEAST STANDARD; PRT; 142 AA.  
 AC Q03525;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 16.2 KD PROTEIN IN PRP24-RRN9 INTERGENIC REGION.  
 GN YMR269W OR YMR156.11.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 CC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA LYE G., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A.;  
 RA SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z49260; G809092; -  
 CC HYPOTHETICAL PROTEIN.  
 FT DOMAIN 61 64 POLY-ARG.  
 FT DOMAIN 89 98 POLY-LYS.  
 SQ SEQUENCE 142 AA; 16155 MW; 747C7BF8 CRC32;

Query Match 90.0%; Score 45; DB 1; Length 142;  
 Best Local Similarity 90.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10  
 Db 89 KKKKKKKKK 98

## RESULT 14

DKC1\_HUMAN STANDARD; PRT; 514 AA.  
 AC O60832; O43845;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DYRKERIN (NUCLEOLAR PROTEIN NAP57) (CBF5 HOMOLOG).  
 GN DKC1.  
 OS HOMO SAPIENS (HUMAN).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A. AND VARIANTS DKC.  
 RX MEDLINE; 98250167.  
 RA HEISS N.S., KNIGHT S.W., VULLIAMY T.J., KLAUCK S.M., WIEMANN S.,  
 RA MASON P.J., POUSTKA A., DOKAL I.;  
 RT "X-linked dyskeratosis congenita is caused by mutations in a highly  
 RT conserved gene with putative nucleolar functions.";  
 RT

RL NAT. GENET. 19:32-38(1998).  
 RP SEQUENCE FROM N.A.  
 RA JIANG W., CLIFFORD J., KOLTIN Y.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR (BY SIMILARITY).  
 CC -1- DISEASE: DEFECTS IN DKC1 ARE THE CAUSE OF X-LINKED RECESSIVE  
 CC DYSERATOPOSIS CONGENITA (DKC); A DISEASE CHARACTERIZED BY THE EARLY  
 CC MANIFESTATION OF RETICULATE SKIN PIGMENTATION, NAIL DYSTROPHY, AND  
 CC MUCOSAL LEUKOPLAKIA. PROGRESSIVE BONE MARROW FAILURE OCCURS IN  
 CC OVER 80% OF CASES AND IS THE MAIN CAUSE OF EARLY MORTALITY.  
 CC -1- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ224481; E1292822; -  
 DR EMBL: U59151; G2737894; -  
 DR NM: 300126; -  
 DR MM: 305000; -  
 DR KW NUCLEAR PROTEIN; DISEASE MUTATION.  
 FT DOMAIN 11 17 POLY-LYS.  
 FT DOMAIN 498 507 POLY-LYS.  
 FT VARIANT 36 36 F -> V (IN DKC).  
 FT VARIANT 37 37 MISSING (IN DKC).  
 FT VARIANT 40 40 P -> R (IN DKC).  
 FT VARIANT 72 72 L -> Y (IN DKC).  
 FT VARIANT 402 402 G -> E (IN DKC).  
 FT CONFLICT 37 37 L -> F (IN REF. 2).  
 SQ SEQUENCE 514 AA; 57674 MW; 152B672C CRC32;

Query Match 88.0%; Score 44; DB 1; Length 514;  
 Best Local Similarity 90.0%; Pred. No. 8.5;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 DB 498 KKKKKKKK 507  
 |||||

RESULT 15  
 IF2P\_HUMAN  
 ID IF2P\_HUMAN STANDARD; PRT; 1220 AA.  
 AC O60841;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PROBABLE TRANSLATION INITIATION FACTOR IF-2.  
 GN KIAA0741.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE; 99087487.  
 RX NAGASE T., ISHIKAWA K.-I., SUYAMA M., KIKUNO R., MIYAJIMA N.,  
 RA TANAKA A., KOTANI H., NOMURA N., OHARA O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XI.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA RES. 5:277-286(1998).  
 RN [2]  
 RP SEQUENCE OF 833-1220 FROM N.A.  
 RA STANCHI F., BERTOCCO E., LANFRANCHI G., VALLE G.;  
 RT "Finding homologues between human and yeast."  
 RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING  
 CC THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO  
 CC FUNCTION ALONG WITH EIF-2 (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB018284; D1035442; -  
 DR EMBL: AJ006412; E1294581; -  
 KW INITIATION FACTOR; PROTEIN BIOSYNTHESIS; GTP-BINDING.  
 FT NP\_BIND 638 645 GTP (POTENTIAL).  
 SQ SEQUENCE 1220 AA; 138754 MW; 48E870FC CRC32;

Query Match 88.0%; Score 44; DB 1; Length 1220;  
 Best Local Similarity 90.0%; Pred. No. 18;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 DB 313 KKKKKKKK 322  
 |||||

Search completed: September 7, 1999, 23:49:57  
 Job time: 1936 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:25:12 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-1  
Perfect score: 50  
Sequence: 1 KKKKKKKKK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	130	2 S55141	hypothetical prote
2	50	100.0	730	2 S34625	probable membrane
3	50	100.0	418	2 S41044	chromosomal protei
4	50	100.0	215	2 I32523	nucleoporin p62 ho
5	47	94.0	478	1 DPHUM2	methionyl aminopep
6	47	94.0	480	2 A46702	probable receptor
7	47	94.0	683	2 T00872	hypothetical prote
8	47	94.0	144	2 S46716	hypothetical prote
9	47	94.0	483	2 F71619	hypothetical prote
10	47	94.0	972	2 F71608	hypothetical prote
11	46	92.0	1647	2 S45252	SNF2beta protein -
12	46	92.0	106	2 S59536	cold stress protei
13	46	92.0	295	2 S61039	hypothetical prote
14	46	92.0	167	2 S38112	hypothetical prote
15	46	92.0	441	2 A48455	acidic phosphoprot
16	46	92.0	2010	2 B71616	phosphatase (acid
17	46	92.0	1138	2 E71615	probable amine tra
18	46	92.0	2013	2 C71610	probable membrane
19	46	92.0	1213	2 A54063	TATA-binding prote
20	46	92.0	404	2 B71224	hypothetical prote
21	45	90.0	437	2 I31238	translation elonga
22	45	90.0	1058	2 T00414	COPI-interacting p
23	45	90.0	961	2 T01167	hypothetical prote
24	45	90.0	1051	2 J04091	glycoprotein A - p
25	45	90.0	142	2 S54481	hypothetical prote
26	45	90.0	558	2 F71609	hypothetical prote
27	45	90.0	1516	2 E71619	RAD2 endonuclease
28	45	90.0	886	2 I38182	transcription term
29	44	88.0	246	2 S49770	hypothetical prote
30	44	88.0	432	2 G71621	protein of the MAK
31	44	88.0	695	2 S40923	hypothetical prote
32	43	86.0	745	2 S49155	heat-shock protein
33	43	86.0	1684	2 T02367	hypothetical prote
34	43	86.0	300	2 T02452	probable glycoprot
35	43	86.0	216	2 S50766	dehydrin-related p
36	43	86.0	1712	2 C71618	hypothetical prote
37	43	86.0	1193	2 G71605	hypothetical prote
38	43	86.0	671	2 A40692	signal recognition
39	42	84.0	1042	2 I50099	H,K-ATPase - giant

40	42	84.0	686	2 A44842	CGMP-gated ion cha
41	42	84.0	690	2 S07103	CGMP-gated ion cha
42	42	84.0	690	2 A42161	CGMP-gated cation
43	42	84.0	688	2 B42161	CGMP-gated cation
44	42	84.0	133	2 I37382	CGMP-gated cation
45	42	84.0	691	2 JC6509	rod cyclic nucleot

ALIGNMENTS

RESULT 1  
S55141  
Hypothetical protein YNL143c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein JTE130; hypothetical protein N1206; hypotheti  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 12-Dec-1997  
C:Accession: S55141; S59246; S63088  
R:Mallet, L.; Bussiereau, F.; Jacquet, M.  
submitted to the EMBL Data Library, November 1994  
A:Description: A 43.5 kb fragment of the chromosome XIV.  
A:Reference number: S55136  
A:Accession: S55141  
A:Molecule type: DNA  
A:Residues: 1-130 <MAL>  
A:Cross-references: EMBL:Z46843; NID:9861113; PID:9854495  
R:Mallet, L.; Bussiereau, F.; Jacquet, M.  
Yeast 11, 1195-1209, 1995  
A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MFA2, MEP2, CAP/SR  
A:Reference number: S59241; MUID:96109932  
A:Accession: S59246  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-130 <MAW>  
A:Cross-references: EMBL:Z46843; NID:9861113; PID:9854495  
R:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63069  
A:Accession: S63088  
A:Molecule type: DNA  
A:Residues: 1-130 <MAF>  
A:Cross-references: EMBL:Z71419; NID:91302092; PID:e233951; PID:g1302093; MIPS:YNL143  
A:Experimental source: strain S288C  
C:Genetics:  
A:Map position: 14L  
Query Match 100.0%; Score 50; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 0.99;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKKKKKKKK 10  
DB 49 KKKKKKKKK 58  
RESULT 2  
S34625  
probable membrane protein YOR019w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein Q2631; hypothetical protein YOL303.8  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 12-Dec-1997  
C:Accession: S34625; S66885  
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54617  
A:Accession: S54625  
A:Molecule type: DNA  
A:Residues: 1-730 <DEH>  
A:Cross-references: EMBL:X87331; NID:g1041652; PID:g829129  
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.  
submitted to the Protein Sequence Database, July 1996

A:Reference number: S66877  
A:Accession: S66885  
A:Molecule type: DNA  
A:Residues: 1-730 <DEW>  
A:Cross-references: EMBL:274927; NID:g1420122; PID:g252323; PID:g1420123; MIPS:YOR019w  
A:Experimental source: strain S288C  
C:Genetics:  
A:Map position: 15R  
C:Keywords: transmembrane protein  
F:317-333/Domain: transmembrane #status predicted <TMM>

Query Match 100.0%; Score 50; DB 2; Length 730;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 713 KKKKKKKKK 722

RESULT 3  
S41044  
chromosomal protein - human  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 07-Feb-1997  
C:Accession: S41044  
R:Yeo, J.P.; Alderuccio, F.; Toh, B.H.  
Nature 367, 288-291, 1994  
A:Title: A new chromosomal protein essential for mitotic spindle assembly.  
A:Reference number: S41044; MUID:94166884  
A:Accession: S41044  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-418 <YEO>  
C:Keywords: chromosomal protein

Query Match 100.0%; Score 50; DB 2; Length 418;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 269 KKKKKKKKK 278

RESULT 4  
I52523  
nucleoporin p62 homolog - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 27-Feb-1997  
C:Accession: I52523  
R:Wang, Z.Q.; Akmal, K.M.; Kim, K.H.  
Biol. Reprod. 51, 1022-1030, 1994  
A:Title: An unusual nucleoporin-related messenger ribonucleic acid is present in the germ  
A:Reference number: I52523; MUID:95151924  
A:Accession: I52523  
A:Status: preliminary; translated from GB/EMBL/DDJB  
A:Molecule type: mRNA  
A:Residues: 1-215 <RES>  
A:Cross-references: GB:S75997; NID:g913245; PID:g913246  
A:Experimental source: testis

RESULT 5  
DPHUM2  
methionyl aminopeptidase (EC 3.4.11.18) 2 - human  
N:Alternate names: p67; translation initiation factor eIF-2-associated protein  
C:Species: Homo sapiens (man)  
C:Date: 14-Jul-1995 #sequence\_revision 23-Aug-1996 #text\_change 26-Feb-1999  
C:Accession: S52112  
R:Li, X.; Chang, Y.H.  
Biochim. Biophys. Acta 1260, 333-336, 1995  
A:Title: Molecular cloning of a human complementary DNA encoding an initiation factor  
A:Reference number: S52112; MUID:95178556  
A:Accession: S52112  
A:Molecule type: mRNA  
A:Residues: 1-478 <LIX>  
A:Cross-references: GB:U13261; NID:g687242; PID:g687243  
C:Genetics:  
A:Gene: GDB:P67EIF2  
A:Cross-references: GDB:512821  
C:Function:  
A:Description: catalyzes hydrolysis of amino-terminal methionine from proteins  
C:Superfamily: human methionyl aminopeptidase  
C:Keywords: alpha-aminoacylpeptide hydrolase; cobalt; metalloprotein; protein biosynt  
F:251,262,459/Binding site: cobalt 2 (Asp, Asp, Glu) #status predicted  
F:262,331,364,459/Binding site: cobalt 1 (Asp, His, Glu, Glu) #status predicted

Query Match 94.0%; Score 47; DB 1; Length 478;  
Best Local Similarity 90.0%; Pred. No. 6.2;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 98 KKKKKKKKKR 107

RESULT 6  
A46702  
methionyl aminopeptidase (EC 3.4.11.18) 2 - rat  
N:Alternate names: p67; translation initiation factor eEF2-associated protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Sep-1997  
C:Accession: A46702  
R:Wu, S.; Gupta, S.; Chatterjee, N.; Hileman, R.E.; Kinzy, T.G.; Denslow, N.D.; Merri  
J. Biol. Chem. 268, 10796-10801, 1993  
A:Title: Cloning and characterization of complementary DNA encoding the eukaryotic in  
A:Reference number: A46702  
A:Accession: A46702  
A:Molecule type: mRNA  
A:Residues: 1-480 <WUA>  
A:Cross-references: GB:L10652; NID:g204003; PID:g204004  
C:Superfamily: human methionyl aminopeptidase  
C:Keywords: alpha-aminoacylpeptide hydrolase; glycoprotein

Query Match 94.0%; Score 47; DB 2; Length 480;  
Best Local Similarity 90.0%; Pred. No. 6.2;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKK 10  
DB 98 KKKKKKKKKR 107

RESULT 7  
T00872  
probable receptor protein kinase Fl7K2.12 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
C:Accession: T00872  
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke  
submitted to the EMBL Data Library, March 1998  
A:Description: Arabidopsis thaliana chromosome II BAC Fl7K2 genomic sequence.

A:Reference number: Z14207  
A:Accession: T00872  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-683 <R0U>  
A:Cross-references: EMBL:AC003680; NID:g2979540; PID:g2979551  
C:Genetics:  
A:Map position: II  
A:Note: F17K3.12

Query Match 94.0%; Score 47; DB 2; Length 683;  
Best Local Similarity 90.0%; Pred. No. 8.1;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
||:|||||||

DB 451 KKKKKKKKK 460

## RESULT 8

S46716  
hypothetical protein YHR095w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein H9332.2  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-Dec-1997  
C:Accession: S46716  
R:Vaudin, M.  
A:Description: The sequence of S. cerevisiae cosmid 9332.  
A:Reference number: S46715  
A:Accession: S46716  
A:Molecule type: DNA  
A:Residues: 1-144 <VAU>  
A:Cross-references: EMBL:U00060; NID:g487928; PID:g487930; MIPS:YHR095w  
C:Genetics:  
A:Map position: 8R

Query Match 94.0%; Score 47; DB 2; Length 144;  
Best Local Similarity 90.0%; Pred. No. 2.5;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
|||||||

DB 111 KKKKKKKKK 120

## RESULT 9

F71619  
hypothetical protein PFB0235w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: F71619  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, E.V.;  
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: F71619  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-483 <GAR>  
A:Cross-references: GB:AE001382; GB:AE001362; NID:g3845130; PID:g3845131; TIGR:PFB0235w  
C:Genetics:  
A:Experimental source: clone 3D7  
A:Gene: PFB0235w

Query Match 94.0%; Score 47; DB 2; Length 483;  
Best Local Similarity 90.0%; Pred. No. 6.2;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
|||||||

DB 449 KKKKKKKKK 458

## RESULT 10

F71608  
hypothetical protein PFB0700c - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: F71608  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, E.;  
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: F71608  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-972 <GAR>  
A:Cross-references: GB:AE001412; GB:AE001362; NID:g3845251; PID:g3845254; TIGR:PFB070  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0700c

Query Match 94.0%; Score 47; DB 2; Length 972;  
Best Local Similarity 90.0%; Pred. No. 11;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
|||||||

DB 567 KKKKKKKKK 576

## RESULT 11

S45252  
SNF2beta protein - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Dec-1994 #sequence\_revision 17-Nov-1995 #text\_change 12-Sep-1997  
C:Accession: S45252  
R:Chiba, H.; Muramatsu, M.; Nomoto, A.; Kato, H.  
Nucleic Acids Res. 22, 1815-1820, 1994  
A:Title: Two human homologues of Saccharomyces cerevisiae SWI2/SNF2 and Drosophila br  
A:Reference number: S45251; MUID:94268902  
A:Accession: S45252  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1647 <CHI>  
C:Superfamily: unassigned bromodomain proteins: bromodomain homology  
F:1485-1540/Domain: bromodomain homology <BRO>

Query Match 92.0%; Score 46; DB 2; Length 1647;  
Best Local Similarity 90.0%; Pred. No. 21;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
|||||||

DB 578 KKKKKKKKK 587

## RESULT 12

S59536  
cold stress protein COR11 - Poncirus trifoliata  
C:Species: Poncirus trifoliata  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Sep-1998  
C:Accession: S59536  
R:Cal, Q.; Moore, G.A.; Guy, C.L.  
Plant Mol. Biol. 29, 11-23, 1995  
A:Title: An unusual group 2 LEA gene family in citrus responsive to low temperature.  
A:Reference number: S59534; MUID:96017610  
A:Accession: S59536

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-106 <CAI>

A:Cross-references: EMBL:L39005; NID:g625154; PID:g625155

A:Experimental source: seedling; tissue type leaf

C:Superfamily: cold stress protein COR19

C:Keywords: cold shock

Query Match 92.0%; Score 46; DB 2; Length 106;

Best Local Similarity 90.0%; Pred. No. 2.6;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10

|||||

Db 83 KKKKKKKKK 92

RESULT 13

S61039

hypothetical protein YDL173w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D1438

C:Species: Saccharomyces cerevisiae

C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 12-Dec-1997

C:Accession: S61039; S67725

R:Pohl, T.M.

submitted to the EMBL Data Library, November 1995

A:Reference number: S61010

A:Accession: S61039

A:Molecule type: DNA

A:Residues: 1-295 <POH>

A:Cross-references: EMBL:Z67750; NID:g1061256; PID:g1061265

R:Pohl, T.M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67708

A:Accession: S67725

A:Molecule type: DNA

A:Residues: 1-295 <POW>

A:Cross-references: EMBL:Z74221; NID:g1431275; PID:e253081; PID:g1431276; MIPS:YDL173w

A:Experimental source: strain S288C

C:Genetics:

A:Map position: 4L

Query Match 92.0%; Score 46; DB 2; Length 295;

Best Local Similarity 90.0%; Pred. No. 5.7;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10

|||||

Db 274 KKKKKKKKK 283

RESULT 14

S38112

hypothetical protein YKR040c - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 14-Nov-1997

C:Accession: S38112

R:Urrstarazu, L.A.; Jauniaux, J.C.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S38097

A:Accession: S38112

A:Molecule type: DNA

A:Residues: 1-167 <URP>

A:Cross-references: EMBL:Z28265; NID:g486480; PID:g486481; MIPS:YKR040c

A:Experimental source: strain S288C

C:Genetics:

A:Map position: 11R

Query Match 92.0%; Score 46; DB 2; Length 167;

Best Local Similarity 90.0%; Pred. No. 3.7;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10

|||||

Db 143 KKKKKKKKK 152

RESULT 15

A48455

acidic phosphoprotein PCBMAlq - Plasmodium chabaudi

C:Species: Plasmodium chabaudi

C:Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998

C:Accession: A48455

R:Deleersnijder, W.; Prasomsitti, P.; Tungpradubkul, S.; Hendrix, D.; Hamers-Casterma

Mol. Biochem. Parasitol. 56, 59-68, 1992

A:Title: Structure of a plasmodium chabaudi acidic phosphoprotein that is associated

A:Reference number: A48455; MUID:93116806

A:Accession: A48455

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-441 <DEL>

A:Cross-references: GB:M95789; NID:g160602; PID:g160603

A:Experimental source: IP-PCI/C

A:Note: sequence extracted from NCBI backbone (NCBIN:121415, NCBIP:121416)

C:Keywords: phosphoprotein

Query Match 92.0%; Score 46; DB 2; Length 441;

Best Local Similarity 90.0%; Pred. No. 7.7;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKK 10

|||||

Db 394 KKKKKKKKK 403

Search completed: September 7, 1999, 23:06:12

Job time: 2460 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:26:36 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-1  
Perfect score: 50  
Sequence: 1 KKKKKKKKKK 10

Scoring table: BLOSUM62

Searched: 105577 seqs, 9858381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	11	1	US-07-694-983-15
2	50	100.0	36	1	US-07-948-357-4
3	50	100.0	36	1	US-07-948-357-12
4	50	100.0	36	1	US-07-948-357-13
5	50	100.0	10	1	US-08-097-830E-1
6	50	100.0	30	1	US-08-097-830E-2
7	50	100.0	434	1	US-08-097-830E-3
8	50	100.0	25	1	US-08-240-514-56
9	50	100.0	398	1	US-08-507-431-2
10	50	100.0	182	2	US-08-226-264-28
11	50	100.0	25	2	US-08-612-302A-56
12	50	100.0	20	2	US-08-769-211-1
13	50	100.0	20	2	US-08-769-211-4
14	50	100.0	10	2	US-08-456-112B-1
15	50	100.0	30	2	US-08-456-112B-2
16	50	100.0	434	2	US-08-456-112B-3
17	50	100.0	398	2	US-08-902-655A-2
18	47	94.0	478	2	US-09-040-793-3
19	46	92.0	1213	1	US-08-188-582-20
20	46	92.0	1213	1	US-08-646-715-20
21	45	90.0	29	1	US-08-054-363-17
22	45	90.0	29	1	US-08-088-658-22
23	45	90.0	29	2	US-08-595-387-17
24	44	88.0	15	2	US-08-701-124-43
25	43	86.0	1497	1	US-08-623-679-7
26	43	86.0	1533	1	US-08-623-679-9
27	41	82.0	1160	3	PCT-US92-05401-4
28	40	80.0	576	1	US-08-190-802A-56
29	40	80.0	12	3	PCT-US95-07543-2
30	39	78.0	32	1	US-08-152-488-8
31	39	78.0	655	1	US-08-264-002-2
32	39	78.0	1087	1	US-08-264-002-5
33	39	78.0	32	1	US-08-303-025-8
34	39	78.0	32	1	US-08-677-304-8
35	39	78.0	18	2	US-08-683-877-2
36	39	78.0	151	2	US-08-872-783-1
37	39	78.0	1612	3	PCT-US94-04496-48
38	38	76.0	501	1	US-08-190-802A-28
39	38	76.0	18	1	US-08-634-060-58

Sequence 18, Appl  
Sequence 1, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 11, Appl  
Sequence 2, Appl

40 38 76.0 732 2 US-08-533-669A-18  
41 38 76.0 953 2 US-08-506-340A-1  
42 38 76.0 501 2 US-08-705-660-10  
43 38 76.0 411 2 US-08-741-134-6  
44 37 74.0 723 1 US-07-814-964-11  
45 37 74.0 1235 1 US-08-118-101A-2

ALIGNMENTS

RESULT 1  
US-07-694-983-15  
; Sequence 15, Application US/07694983  
; Patent No. 5432260  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Philip D.  
; TITLE OF INVENTION: HIGH AFFINITY MANNOSE RECEPTOR  
; TITLE OF INVENTION: LIGANDS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Irell & Manella  
; STREET: 545 Middlefield Road, Suite 200  
; CITY: Menlo Park  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/694,983  
; FILING DATE: 19910503  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 9500-0039.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-327-7250  
; TELEFAX: 415-327-2951  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1  
; OTHER INFORMATION: /label= Ac-  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 11  
; OTHER INFORMATION: /label= -NH2  
US-07-694-983-15

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKK 10  
Db 2 KKKKKKKKKK 11

RESULT 2  
US-07-948-357-4

```
; Sequence 4, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; COMPLEXES INTO HIGHER EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0940004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-07-948-357-4

Query Match 100.0%; Score 50; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10
Db 25 KKKKKKKKK 34

RESULT 3
US-07-948-357-12
; Sequence 12, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; COMPLEXES INTO HIGHER EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0940004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-07-948-357-12

Query Match 100.0%; Score 50; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10
Db 27 KKKKKKKKK 36

RESULT 4
US-07-948-357-13
; Sequence 13, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; COMPLEXES INTO HIGHER EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; CLASSIFICATION: 514
```

ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0940004  
TELEPHONE: (202) 456-0800  
TELEFAX: (202) 833-8716  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
US-07-948-357-13

Query Match 100.0%; Score 50; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 27 KKKKKKKKK 36

RESULT 5  
US-08-097-830E-1  
Sequence 1, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPID A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
PRIOR APPLICATION DATA:  
CLASSIFICATION: 514  
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-097-830E-1  
Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 1 KKKKKKKKK 10

RESULT 6  
US-08-097-830E-2  
Sequence 2, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPID A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-097-830E-2  
Query Match 100.0%; Score 50; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
DB 1 KKKKKKKKK 10

RESULT 7  
US-08-097-830E-3  
Sequence 3, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPID A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,659  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8998  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-3

Query Match 100.0%; Score 50; DB 1; Length 434;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
Db 1 KKKKKKKKK 10

RESULT 8  
US-08-240-514-56  
Sequence 56, Application US/08240514  
Patent No. 5670347  
GENERAL INFORMATION:  
APPLICANT: GOPAL, T. Venkat  
TITLE OF INVENTION: PEPTIDE-MEDIATED GENE TRANSFER  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240,514  
FILING DATE: 11-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 73521/102/CLIN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
US-08-240-514-56

Query Match 100.0%; Score 50; DB 1; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
Db 14 KKKKKKKKK 23

RESULT 9  
US-08-507-431-2  
Sequence 2, Application US/08507431  
Patent No. 5693518  
GENERAL INFORMATION:  
APPLICANT: Kofod, Lene V.  
APPLICANT: Kauppinen, Markus S.  
APPLICANT: Christgau, Stephan  
APPLICANT: Heldt-Hansen, Hans P.  
APPLICANT: Dalboge, Henrik  
APPLICANT: Andersen, Lene N.  
APPLICANT: Si, Joan Q.  
APPLICANT: Jacobson, Tina  
APPLICANT: Munk, Niels  
APPLICANT: Mullertz, Anette  
TITLE OF INVENTION: ENZYMES WITH XYLANASE ACTIVITY FROM  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5693518o No. 5693518disk of No. 5693518th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/507,431  
FILING DATE: 15-FEB-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/002,800  
FILING DATE: 25-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Harrington, James J.  
REGISTRATION NUMBER: 38,711  
REFERENCE/DOCKET NUMBER: 3954.204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-507-431-2

Query Match 100.0%; Score 50; DB 1; Length 398;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
Db 389 KKKKKKKKK 398

RESULT 10  
US-08-226-264-28  
Sequence 28, Application US/08226264  
Patent No. 5801017  
GENERAL INFORMATION:  
APPLICANT: Werber, Moshe M.  
APPLICANT: Zeelon, Elisha P.  
APPLICANT: Levanon, Avigdor  
APPLICANT: Guy, Rachel  
APPLICANT: Goldlust, Arie  
APPLICANT: Rigbi, Meir  
APPLICANT: Panet, Amos  
APPLICANT: Fischer, Meir  
TITLE OF INVENTION: PRODUCTION OF RECOMBINANT FACTOR XA  
TITLE OF INVENTION: INHIBITORS OF LEECH HIRUDO MEDICINALIS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/226,264  
FILING DATE: 08-APR-94  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 40017-A/JPW/GJG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-378-0400  
TELEFAX: 212-391-0525  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 182 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-226-264-28  
Query Match 100.0%; Score 50; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKKKKKKK 10  
Db 168 KKKKKKKK 177  
RESULT 11  
US-08-612-302A-56  
Sequence 56, Application US/08612302A  
Patent No. 5811297  
Patent No. 5811297 5780297  
GENERAL INFORMATION:  
APPLICANT: GOPAL, T. Venkat  
TITLE OF INVENTION: PEPTIDE-MEDIATED GENE TRANSFER  
NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: One Westlakes, Berwyn  
CITY: Valley Forge  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/612,302A  
FILING DATE: 7 March 1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Amzel, Viviana  
REGISTRATION NUMBER: 30,930  
REFERENCE/DOCKET NUMBER: AMBA-020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 407-0700  
TELEFAX: (610) 407-0701  
TELEX: n.a.  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: n.a.  
TOPOLOGY: n.a.  
US-08-612-302A-56  
Query Match 100.0%; Score 50; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKKKKKKK 10  
Db 14 KKKKKKKK 23  
RESULT 12  
US-08-769-211-1  
Sequence 1, Application US/08769211  
Patent No. 5830852  
GENERAL INFORMATION:  
APPLICANT: Thatcher et al.  
TITLE OF INVENTION: Improved Pharmaceutical Compositions  
TITLE OF INVENTION: Insulin-Receptor Mediated Gene Therapy.  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Inc.  
STREET: 75 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1807  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/769,211  
FILING DATE: 18-Dec-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kathleen M. Williams  
REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3255/02514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-345-9100  
TELEFAX: 617-345-9111  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: Xaa at position 20 represents  
OTHER INFORMATION: S-acetimidomethyl-Cys or S-pyridyl-Cys.  
US-08-769-211-1

Query Match 100.0%; Score 50; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.24;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10  
DB 2 KKKKKKKK 11

RESULT 13  
US-08-769-211-4  
Sequence 4, Application US/08769211  
Patent No. 5830852  
GENERAL INFORMATION:  
APPLICANT: Thatcher et al.  
TITLE OF INVENTION: Improved Pharmaceutical Compositions  
TITLE OF INVENTION: Insulin-Receptor Mediated Gene Therapy.  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Inc.  
STREET: 75 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1807

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/769,211  
FILING DATE: 18-Dec-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kathleen M. Williams  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: 3255/02514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-345-9100  
TELEFAX: 617-345-9111  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-769-211-4

Query Match 100.0%; Score 50; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.24;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10  
DB 2 KKKKKKKK 11

RESULT 14  
US-08-456-112B-1  
Sequence 1, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8988  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-1

Query Match 100.0%; Score 50; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKK 10  
DB 1 KKKKKKKK 10

RESULT 15  
US-08-456-112B-2  
Sequence 2, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-2

Query Match 100.0%; Score 50; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KKKKKKKKK 10  
| | | | |  
Db 1 KKKKKKKKK 10

Search completed: September 7, 1999, 22:38:29  
Job time: 7913 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 15:29:31 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-1  
Perfect score: 50  
Sequence: 1 KKKKKKKKKK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	175	1 R08261	Peptide antigenic
2	50	100.0	40	1 R11512	Anti-ATLA antibody
3	50	100.0	29	1 R11513	Anti-ATLA antibody
4	50	100.0	34	1 R11514	Anti-ATLA antibody
5	50	100.0	898	1 P61030	Entire coded sequ
6	50	100.0	899	1 P61036	Translation of pla
7	50	100.0	898	1 P61082	Complete translati
8	50	100.0	20	1 P20159	Sequence of lysine
9	50	100.0	36	1 R34595	Sequence of endoso
10	50	100.0	36	1 R34603	Sequence of endoso
11	50	100.0	36	1 R34604	Sequence of endoso
12	50	100.0	412	1 R63789	Aspergillus aculea
13	50	100.0	182	1 R62619	Deduced sequence o
14	50	100.0	418	1 R63676	Mitosis-associated
15	50	100.0	16	1 R65867	Antideoxyribonucle
16	50	100.0	17	1 R66010	Peptide which bind
17	50	100.0	10	1 R71774	Peptide neutralisi
18	50	100.0	25	1 R88385	Gene therapy synth
19	50	100.0	19	1 R94325	Anti-thrombogenic
20	50	100.0	20	1 R94326	Anti-thrombogenic
21	50	100.0	116	1 W03642	Human cannabinoid
22	50	100.0	36	1 W01422	8toxin deriv, endo
23	50	100.0	36	1 W01430	Melittin peptide 1
24	50	100.0	36	1 W01431	Melittin peptide 2
25	50	100.0	10	1 W21589	Antibiotic potenti
26	50	100.0	30	1 W21590	Antibiotic potenti
27	50	100.0	434	1 W21591	Antibiotic potenti
28	50	100.0	20	1 W08403	Nucleic acid conde
29	50	100.0	17	1 W24448	Nucleic acid (NA)
30	50	100.0	45	1 W24450	Nucleic acid (NA)
31	50	100.0	14	1 W24446	Nucleic acid (NA)
32	50	100.0	15	1 W24447	Nucleic acid (NA)
33	50	100.0	19	1 W32265	Peptide used in ge
34	50	100.0	20	1 W32266	Peptide used in ge
35	50	100.0	40	1 W24855	Bifunctional pepti
36	50	100.0	19	1 W38873	Delivery peptide u
37	50	100.0	28	1 W38839	Delivery peptide u
38	50	100.0	20	1 W38874	Delivery peptide u
39	50	100.0	29	1 W38840	Delivery peptide u
40	50	100.0	21	1 W38875	Delivery peptide u
41	50	100.0	30	1 W38841	Delivery peptide u
42	50	100.0	22	1 W38876	Delivery peptide u
43	50	100.0	31	1 W38842	Delivery peptide u

#### ALIGNMENTS

```

RESULT 1
R08261 ID R08261 standard; protein; 175 AA.
AC R08261;
DT 18-FEB-1991 (first entry)
DE Peptide antigenic for malarial antisporezoite Antibodies.
KW Malaria; vaccine; universal carrier molecule.
OS Synthetic.
FH Key Location/Qualifiers
FT duplication 1..160
FT /label= (-Asn-Ala-Ala-Gly-)3-40
FT duplication 161..175
FT /label= (-Lys-)1-15 (odd number)
FT cross_links 161..175
FT /label= nth Lys has side chain identical to
FT AAs 1..n-1
PN EP-398443-A.
PD 22-NOV-1990.
PF 16-MAY-1990; 201244.
PR 19-MAY-1989; IT-020553.
PR 23-MAR-1990; IT-019800.
PA (ENIE ) ENRICERCH SPA.
PI Pessi A, Bonelli F, Chiappinelli L, Bianchi E, Del Giudice G;
DR WPI; 90-350221/47.
DR New immunogenic cpds. for synthetic vaccines for Plasmodium
PT malariae - and for determ. of antibodies of Plasmodium malariae
PS Claim 1; Page 15; 19pp; English.
CC Peptides have a fractal structure, branching at each lysine residue
CC into a side chain identical to all AAs at its N-terminal.
CC Peptides are useful as universal carrier molecules inducing
CC genetically unrestricted protective immunity against different
CC pathogenic agents.
CC Specifically, the peptides may be used as antigens in vaccination
CC and immunoassay of Plasmodium malariae antisporezoite antibodies.
CC See also R08260.
SQ Sequence 175 AA;

Query Match 100.0%; Score 50; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 0.88; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

Qy 1 KKKKKKKKKK 10
    |||||
Db 161 KKKKKKKKKK 170

RESULT 2
R11512 ID R11512 standard; peptide; 40 AA.
AC R11512;
DT 21-JUN-1991 (first entry)
DE Anti-ATLA antibody-binding peptide #1.
KW adult T cell leukaemia associated antigen; HTLV-1; ATL.
OS Synthetic.
PN EP-423649-A.
PD 24-APR-1991.
PF 12-OCT-1990; 119624.
PR 13-OCT-1989; JP-266983.
PA (KURS ) KURARAY KK.
PI Maeda Y, Shiraki H, Washitani Y, Kuroda N, Yamada K;
PI Oka K, Namba T;
DR WPI; 91-118942/17.
PT New peptide(s) able to bind to anti-ATLA antibody - used in
PT diagnosis and to treat HTLV-1-associated myelopathy and diseases
PT caused by ATL

```

44 50 100.0 23 1 W38877 Delivery peptide u  
45 50 100.0 32 1 W38843 Delivery peptide u

PS Claim 2; Page 31; 44 pp; English.  
 CC This peptide is able to bind to an antibody specific against an  
 CC adult T cell leukaemia associated antigen. The peptide can have  
 CC from 1 to 10 Lys residues at the N-terminus. The invention also  
 CC covers an adsorbent for anti-ATLA antibodies, comprising the peptide  
 CC immobilised on a carrier and a reagent containing the peptide, for  
 CC measuring the antibody. The reagent and the adsorbent can be used  
 CC to treat an adult T-cell leukaemia virus-infectious disease.  
 CC See also R11513 and R11514.  
 SQ Sequence 40 AA;

Query Match 100.0%; Score 50; DB 1; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 0.29;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 DB 1 KKKKKKKKK 10

RESULT 3  
 R11513  
 ID R11513 standard; Protein; 29 AA.  
 AC R11513;  
 DT 21-JUN-1991 (first entry)  
 DE Anti-ATLA antibody-binding peptide #2.  
 KW adult T cell leukaemia associated antigen; HTLV-1; ATL.  
 OS Synthetic.  
 PN EP-423649-A.  
 PD 24-APR-1991.  
 PF 12-OCT-1990; 119624.  
 PR 13-OCT-1989; JP-266983.  
 PA (KURS ) KURARAY KK.  
 PI Maeda Y, Shiraki H, Washitani Y, Kuroda N, Yamada K;  
 DR WPI; 91-118942/17.  
 PT New peptide(s) able to bind to anti-ATLA antibody - used in  
 PT diagnosis and to treat HTLV-1-associated myelopathy and diseases  
 PT caused by ATL

PS Claim 3; Page 31; 44 pp; English.  
 CC This peptide is able to bind to an antibody specific against an  
 CC adult T cell leukaemia associated antigen. The peptide can have  
 CC from 1 to 10 Lys residues at the N-terminus. The invention also  
 CC covers an adsorbent for anti-ATLA antibodies, comprising the peptide  
 CC immobilised on a carrier and a reagent containing the peptide, for  
 CC measuring the antibody. The reagent and the adsorbent can be used  
 CC to treat an adult T-cell leukaemia virus-infectious disease.  
 CC See also R11512 and R11514.  
 SQ Sequence 29 AA;

Query Match 100.0%; Score 50; DB 1; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.23;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 DB 1 KKKKKKKKK 10

RESULT 4  
 R11514  
 ID R11514 standard; Protein; 34 AA.  
 AC R11514;  
 DT 21-JUN-1991 (first entry)  
 DE Anti-ATLA antibody-binding peptide #3.  
 KW adult T cell leukaemia associated antigen; HTLV-1; ATL.  
 OS Synthetic.  
 PN EP-423649-A.  
 PD 24-APR-1991.  
 PF 12-OCT-1990; 119624.  
 PR 13-OCT-1989; JP-266983.

PA (KURS ) KURARAY KK.  
 PI Maeda Y, Shiraki H, Washitani Y, Kuroda N, Yamada K;  
 DR WPI; 91-118942/17.  
 PT New peptide(s) able to bind to anti-ATLA antibody - used in  
 PT diagnosis and to treat HTLV-1-associated myelopathy and diseases  
 PT caused by ATL  
 PS Claim 4; Page 31; 44 pp; English.  
 CC This peptide is able to bind to an antibody specific against an  
 CC adult T cell leukaemia associated antigen. The peptide can have  
 CC from 1 to 10 Lys residues at the N-terminus. The invention also  
 CC covers an adsorbent for anti-ATLA antibodies, comprising the peptide  
 CC immobilised on a carrier and a reagent containing the peptide, for  
 CC measuring the antibody. The reagent and the adsorbent can be used  
 CC to treat an adult T-cell leukaemia virus-infectious disease.  
 CC See also R11512 and R11513.  
 SQ Sequence 34 AA;

Query Match 100.0%; Score 50; DB 1; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 DB 1 KKKKKKKKK 10

RESULT 5  
 P61030  
 ID P61030 standard; Protein; 898 AA.  
 AC P61030;  
 DT 13-AUG-1991 (first entry)  
 DE Entire coded sequence from plasmid pAU157 insert encoding rat-liver  
 DE P-450MC cytochrome.  
 OS 3-methyl-cholethrene; MC; ds.  
 KW Rattus sp.  
 PN J61052284-A.  
 PD 14-MAR-1986.  
 PF 15-AUG-1984; 169447.  
 PR 15-AUG-1984; JP-169447.  
 PR 24-AUG-1984; JP-175159.  
 PA (AGEN ) AGENCY OF IND SCI TECH.  
 DR WPI; 86-109962/17.  
 DR N-PSDB; N60909.  
 PT New plasmid used in coding rat-liver cytochrome p-450 gene.  
 PS Disclosure; Fig 1; 13pp; Japanese.  
 CC The gene product may be produced in commercial quantities from a  
 CC transformed expression system. The rat-liver cytochrome is readily  
 CC separated, and is induced by 3-methyl-cholethrene.  
 CC AAs marked as x correspond to stop codons in the insert.  
 SQ Sequence 898 AA;

Query Match 100.0%; Score 50; DB 1; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 3;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10  
 |||||  
 DB 873 KKKKKKKKK 882

RESULT 6  
 P61056  
 ID P61056 standard; Protein; 899 AA.  
 AC P61056;  
 DT 08-JUL-1991 (first entry)  
 DE Translation of plasmid pAU157 encoding rat liver cytochrome p-450MC.  
 KW Organic oxide removal; tac-promoter.  
 OS Rattus rattus.  
 PN J61005783-A.  
 PD 11-JAN-1986.

```

PF 16-JUN-1984; 122952.
PR 16-JUN-1984; JP-122952.
PA (AGEN ) AGENCY OF IND SCI TECH.
DR WPI: 86-052223/08.
DR N-PSDB; N60864.
PT Plasmid pmG1 - is used for genetic expression of rat liver
PT cytochrome P-450MC gene in Escherichia coli.
PS Disclosure; Fig 1; 9pp; Japanese.
CC The plasmid may be expressed from a transformed E.coli host, under
CC the control of a strong upstream tac-promoter. The product
CC may be useful in the removal of oxides from organic compounds in
CC industrial waste by an oxidation reaction.
CC In this translation, all stop codons are given as unknown (x).
SQ Sequence 899 AA;

Query Match 100.0%; Score 50; DB 1; Length 899;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKK 10
Db 873 KKKKKKKKKK 882

RESULT 7
PF1082
ID P61082 standard; Protein: 898 AA.
AC P61082;
DT 04-OCT-1991 (first entry)
DE Complete translation of plasmid pAU157 sequence encoding rat liver
DE P-450MC cytochrome.
KW Yeast; alcohol dehydrogenase promoter.
OS Rattus sp.
PN J61088878-A.
PD 07-MAY-1986.
PF 16-JUN-1984; 122953.
PR 16-JUN-1984; JP-122953.
PR 24-AUG-1984; JP-175159.
PA (AGEN ) AGENCY OF IND SCI TECH.
DR WPI: 86-157863/25.
DR N-PSDB; N60889
PT Plasmid for manifestation of cytochrome P-450MC in yeast -
PT contains rats liver cytochrome P-450MC gene.
PS Disclosure; Fig 1; 13pp; Japanese.
CC The cytochrome may be expressed by plasmid pAU157 under the control
CC of an upstream alcohol dehydrogenase promoter and downstream adh
CC terminator. The gene product may be produced from a transformed yeast
CC expression system in large quantities, and may be used for
CC oxidation (fixing with NADPH-cytochrome P-450 reducing enzyme)
CC in a bioreactor or in waste treatment.
CC Residues marked as x correspond to termination sites ignored in the
CC specifications translation.
SQ Sequence 898 AA;

Query Match 100.0%; Score 50; DB 1; Length 898;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKK 10
Db 873 KKKKKKKKKK 882

RESULT 8
PF20159
ID P20159 standard; peptide: 20 AA.
AC P20159;
DT 19-AUG-1992 (first entry)
DE Sequence of lysine polymer.
KW Penicillin allergy; diagnostic; beta-lactam antibiotic; lysine.
FH Key Location/Qualifiers

```

```

FT modified_site 1 /label= H-Lys
FT modified_site 20 /label= Lys-OH
FT misc_difference 1. 20
FT misc_difference 1. 20 /note= "Comprises from 8-20 Lys residues"
PN EP--47197-A.
PD 10-MAR-1982.
PF 06-AUG-1981; 401267.
PR 29-AUG-1980; FR-018809.
PA (DWECK/) DE WECK A L.
PA (IMMO-) INST IMMUNOLOGIE IN.
PA (KLIN-) INST KLINISCHE IMM.
PI Deweck AL, Schneider CH, Rolli HP;
DR WPI: 82-20236E/11 (20236E).
PT Diagnostics for allergic reactions to penicillin(s) - and other
PT beta-lactam antibiotics are homogeneous lysine polymers, contg.
PS Claim 1; Page 12; 15pp; French.
CC The Lys polymers of the invention are used as supports in the prepn.
CC of products for detecting allergies to a penicillin or other beta-
CC lactam type antibiotics. The diagnostic products contain conjugation
CC products of the Lys polymers with benzyl-penicillin or another beta-
CC lactam type antibiotic. The advantage of the claimed diagnostic
CC product over existing products is its homogeneity; it contains 8-12
CC benzoyl penicilloyl joined to a 8-12 Lys molecule skeleton giving
CC optimum conditions for producing allergic erythema reactions.
SQ Sequence 20 AA;

Query Match 100.0%; Score 50; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKK 10
Db 1 KKKKKKKKKK 10

RESULT 9
R34595
ID R34595 standard; peptide: 36 AA.
AC R34595;
DT 09-AUG-1993 (first entry)
DE Sequence of endosomolytic peptide.
KW Endosomolytic peptide; endosome; transfection; gene transfer.
OS Synthetic.
PN W09307283-A.
PD 15-APR-1993.
PF 28-SEP-1992; E02234.
PR 30-SEP-1991; US-767788.
PR 30-SEP-1991; US-768039.
PR 30-JAN-1992; US-827102.
PR 30-JAN-1992; US-827103.
PR 07-APR-1992; US-864759.
PR 02-SEP-1992; US-937788.
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
PA (GETH ) GENENTECH INC.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Birnstiel ML, Cotten M, Curiel D, Oberhauser B,
PI Plank C, Schmidt WGM, Wagner E, Zatloukal K;
DR WPI: 93-134470/16.
PT Compens. for efficient transfection of higher eukaryotic cells -
PT contains conjugate of nucleic acid, and affinity substance plus
PT endosomolytic agent, esp. adenovirus, useful e.g. in gene
PT transfer therapy
PS Example; Page 192; 281pp; German.
CC The peptide is endosomolytic. It can release the contents of
CC endosomes into the cytoplasm. It can be covalently or non-
CC covalently bound to a nucleic acid affinity substance (NAS) directly
CC or via an NA-binding domain (NA = nucleic acid). The compens. are
CC esp. used to introduce NA into human cells.

```

```
SQ      Sequence      36 AA;

Query Match      100.0%; Score 50; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KKKKKKKKK 10
        |||||
DB      25 KKKKKKKKK 34

RESULT 10
R34603
ID      R34603 standard; peptide; 36 AA.
AC      R34603;
DE      09-AUG-1993 (first entry)
KW      Sequence of endosomolytic peptide.
OS      Endosomolytic peptide; endosome; transfection; gene transfer.
PN      WO9307283-A.
PD      15-APR-1993.
PF      28-SEP-1992; E02234.
PR      30-SEP-1991; US-767788.
PR      30-SEP-1991; US-768039.
PR      30-JAN-1992; US-827102.
PR      30-JAN-1992; US-827103.
PR      07-APR-1992; US-864759.
PR      02-SEP-1992; US-937788.
PA      (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
PA      (GETH ) GENENTECH INC.
PA      (UYNC-) UNIV NORTH CAROLINA.
PI      Birnstiel ML, Cotten M, Curiel D, Oberhauser B,
PI      Plank C, Schmidt WGM, Wagner E, Zatloukal K;
DR      WPI; 93-134470/16.
PT      Compsn. for efficient transfection of higher eukaryotic cells -
PT      contains conjugate of nucleic acid, and affinity substance plus
PT      endosomolytic agent, esp. adenovirus, useful e.g. in gene
PT      transfer therapy
PS      Example; Page 196; 281pp; German.
CC      The peptide is endosomolytic. It can release the contents of
CC      endosomes into the cytoplasm. It can be covalently or non-
CC      covalently bound to a nucleic acid affinity substance (NAS) directly
CC      or via an NA-binding domain (NA = nucleic acid). The compsns. are
CC      esp. used to introduce NA into human cells.
SQ      Sequence 36 AA;

Query Match      100.0%; Score 50; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KKKKKKKKK 10
        |||||
DB      27 KKKKKKKKK 36

RESULT 12
R63789
ID      R63789 standard; Protein; 412 AA.
AC      R63789;
DE      07-JUN-1995 (first entry)
DE      Aspergillus aculeatus xylanase I.
KW      Xylanase I; Aspergillus aculeatus; brewing; paper pulp;
KW      food preparation; plant cell wall degradation.
OS      Aspergillus aculeatus.
FH      Key Location/Qualifiers
FT      misc_difference 104
FT      /note= "corresponding codon codes Lys"
FT      misc_difference 193
FT      /note= "corresponding nucleotides GC
FT      resultant frame shift"
FT      misc_difference 243
FT      /note= "corresponding codon codes Cys"
FT      misc_difference 404..412
FT      /note= "no corresponding codons"
FT      WO9421785-A.
PN      29-SEP-1994.
PD      02-MAR-1994; DK00088.
PF      10-MAR-1993; DK-000268.
PR      14-OCT-1993; DK-001151.
PA      (NOVO ) NOVO-NORDISK AS.
PA      Andersen LN, Christgau S, Dalboge H, Heldt-hansen HP;
PI      Jacobsents, Kauppinen MS, Kofod LV, Mullertz A;
PI      Munk N, Si JQ;
DR      WPI; 94-317006/39.
DR      N-PSDB; Q74636.
PT      New xylanase enzymes from Aspergillus aculeatus - used for
PT      degrading plant cell wall components, e.g. in the prepn. of feed,
PT      in baking and in prepn. of pulp or paper
PS      Claim 9; Page 57; 80pp; English.
CC      Q74636 encodes R63789 Aspergillus aculeatus xylanase I, which
CC      degrades plant cell wall components and reduces the viscosity
CC      of plant cell wall derived material. These properties are
CC      useful in the production of dough and baked products; in the
CC      preparation of feed, food, beer, wine, pulp and paper; and for the
CC      separation of cereal components. It can also be used in the
CC      production of antibodies.
SQ      Sequence 412 AA;
```

```

Query Match      100.0%; Score 50; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10
    |||||
Db 394 KKKKKKKKK 403

RESULT 13
R62619 ID R62619 standard; Protein; 182 AA.
AC R62619: 1995 (first entry)
DE Deduced sequence of clone pSP65-xai-11 encoding recombinant
KW Factor Xa inhibitor (FXaI).
OS Hirudo medicinalis.
FH Key Location/Qualifiers
FT peptide 1..25
FT protein 26..110
FT peptide 1..110
FT /label= FXaI prepeptide
FT misc_difference 111..182
FT /label= X = posn. of stop codon in Q72756
PN W09423709-A.
PD 27-OCT-1994.
PF 08-APR-1994; U03918.
PR 09-APR-1993; US-045804.
PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
PA (YISS ) YISSUM RES & DEV CO.
PI Fischer M, Goldlust A, Guy R, Levanon A, Panet A;
PI Rigbi M, Werber MM, Zeelon EP;
DR WPI: 94-341457/42.
DR N-PSDB; Q72956.
PT Recombinant factor 10a inhibitor of Hirudo medicinalis - for
PT treating excessive blood coagulation, partic. thrombosis, also
PT related DNA, vectors, transformed cells and antibodies
PS Disclosure; Figure 10; 107pp; English.
CC Total RNA was extracted from 120 leeches and used to prepare a
CC cDNA library which was screened for clones contg. FXaI DNA using
CC as probe radiolabeled DNA from the PCR derived plasmid pSP65-xai-
CC 4 (see Q72957). The plasmid of one of the positive clones - pSK-
CC xai-11 (clone 11) contains a 684 bp insert (see Q72956) essentially
CC identical to the previously obt'd. PCR derived sequences but contg.
CC 78 additional bps at its 5' end. It encodes a prepeptide of 110
CC AAs (incl. the N-terminal Met) extending from Met 1 to Gly 110 (see
CC R62619). AAs 1-25 constitute a leader sequence, and the mature
CC protein comprises Tyr 26 to Gly 110. The sequence Tyr 2 - Gly 86
CC of clone 4 (Q72957) is equiv. to the sequence Tyr 26 - Gly 110
CC of clone 11 (Q72956). There are only two differences between
CC clones 4 and 11. The polypeptide expressed by clone 11 is
CC referred to as recombinant FXaI. A plasmid comprising the DNA
CC designated pSP65-xai-11 and deposited under ATCC Accession No.
CC 69138 is claimed.
SQ Sequence 182 AA;

Query Match      100.0%; Score 50; DB 1; Length 182;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10
    |||||
Db 168 KKKKKKKKK 177

RESULT 14
R63676 ID R63676 standard; Protein; 418 AA.
AC R63676: 1995 (first entry)
DE Mitosis-associated nuclear antigen RMSA-1.
KW RMSA-1; mitosis-associated nuclear antigen.
FH Key Location/Qualifiers
FT misc_difference 269..278
FT /note= "nuclear localization signal"
FT misc_difference 327..332
FT /note= "nuclear localization signal"
FT modified_site 207..208
FT /note= "phosphorylation consensus motif"
FT misc_difference 259..260
FT /note= "phosphorylation consensus motif"
FT misc_difference 395..396
FT /note= "phosphorylation consensus motif"
PN W09423029-A.
PD 13-OCT-1994.
PF 31-MAR-1994; AU0158.
PR 31-MAR-1993; AU-008067.
PA (MONU ) UNIV MONASH.
PI Alderuccio F, Toh BH, Yeo JP;
DR WPI: 94-333190/41.
DR P-PSDB; R63676.
PT New mitosis-associated nuclear antigen and related DNA - useful
PT for blocking mitosis, e.g. for treatment of cancer, for
PT karyotyping and for diagnosing auto-immune disease
PS Claim 22; Fig 4; 80pp; English.
CC Antisense nucleic acid against the mitosis-associated nuclear
CC antigen RMSA-1 (regulator of mitotic spindle assembly-1) protein
CC can be used to block the activity of this antigen to disrupt/arrest
CC mitosis, especially for the treatment of cancer.
SQ Sequence 418 AA;

Query Match      100.0%; Score 50; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKK 10
    |||||
Db 269 KKKKKKKKK 278

RESULT 15
R65867 ID R65867 standard; peptide; 16 AA.
AC R65867:
DT 26-JUN-1995 (first entry)
DE Antideoxyribonucleic acid antibody adsorbent peptide.
KW Antideoxyribonucleic acid adsorbent peptide; kidney inflammation;
KW systemic lupus erythematosus.
OS Synthetic.
PN J06261941-A.
PD 20-SEP-1994.
PF 13-OCT-1993; 256111.
PR 18-JAN-1993; JP-006100.
PA (KURS ) KURARAY CO LTD.
DR WPI: 94-337463/42.
PT Adsorbents of antideoxyribonucleic - comprising a peptide,
PT immobilised on a carrier.
PS Claim 1; Page 15; 16pp; Japanese.
CC R65841-R65879 are antideoxyribonucleic (ADN) acid adsorbent peptides
CC derived from the highly generic peptide in the specification. When these
CC peptides are immobilised on a carrier (either cellulose, porous
CC glass, polyvinyl alcohol or polyacrylamide) they can be used to
CC remove ADN antibodies, which can result in kidney inflammation and
CC systemic lupus erythematosus.
SQ Sequence 16 AA;

Query Match      100.0%; Score 50; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KKKKKKKKKK 10  
| | | | |  
Db 1 KKKKKKKKKK 10

Search completed: September 7, 1999, 20:36:59  
Job time: 18448 sec



RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA GATTUNG S., SCHEET P., KEMP K.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U80447; AAB37808.1; -;  
 SQ SEQUENCE 560 AA; 65831 MW; 768A2DCE CRC32;

Query Match 72.0%; Score 108; DB 5; Length 560;  
 Best Local Similarity 70.0%; Pred. No. 0.0021;  
 Matches 21; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30  
 |||||:||||:||||:||||:||||:||||  
 Db 270 KKKKKKKKKKKKKKKKKKKKKKK 299

RESULT 3  
 Q21827  
 ID Q21827 PRELIMINARY; PRT; 812 AA.  
 AC Q21827;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE R07E5.1 PROTEIN.  
 GN R07E5.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA AINSCOUGH R.;  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 232863; CAA83621.1; -;  
 SQ SEQUENCE 812 AA; 92630 MW; 9C6E7053 CRC32;

Query Match 72.0%; Score 108; DB 5; Length 812;  
 Best Local Similarity 70.0%; Pred. No. 0.0026;  
 Matches 21; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30

Db 762 KKKKKKKKKKKKKKKKKKKKKKK 791  
 |||||:||||:||||:||||:||||:||||  
 RESULT 4  
 O35807  
 ID O35807 PRELIMINARY; PRT; 129 AA.  
 AC O35807;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE MICROVASCULAR ENDOTHELIAL DIFFERENTIATION GENE 2.  
 GN MDG2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EPIDIDYMAL;  
 RA PROELS F., LOSER B., MARX M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y08769; CAA70022.1; -;  
 DR PFAM; PF00069; pkinase; 1.  
 SQ SEQUENCE 129 AA; 15080 MW; 7A874092 CRC32;

Query Match 71.3%; Score 107; DB 11; Length 129;  
 Best Local Similarity 95.7%; Pred. No. 0.0011;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 23  
 |||||:||||:||||:||||:||||:||||  
 Db 85 KKKKKKKKKKKKKKKKKKKKKKK 107

RESULT 5  
 P97762  
 ID P97762 PRELIMINARY; PRT; 213 AA.  
 AC P97762;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE PAP-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARADA Y., SATO E., IMAMURA Y., KATO H., ARIGA H., IGUCHI-ARIGA S.M.;  
 RT "PAP-1, a novel protein associated with Plm-1.";  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D78255; BAAL1319.1; -;  
 SQ SEQUENCE 213 AA; 25261 MW; AA6B182D CRC32;

Query Match 71.3%; Score 107; DB 11; Length 213;  
 Best Local Similarity 66.7%; Pred. No. 0.0014;  
 Matches 20; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30  
 |||||:||||:||||:||||:||||:||||  
 Db 176 KKKKKKKKKKKKKKKKKKKKKKK 205

RESULT 6  
 O77367  
 ID O77367 PRELIMINARY; PRT; 166 AA.  
 AC O77367;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE MAL3P6.28 PROTEIN.  
 GN MAL3P6.28.



ET	REPEAT	775	839	10	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)
ET	DOMAIN	933	936	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1119	1122	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1202	1205	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1274	1277	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1280	1285	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1287	1294	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1307	1311	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	
ET	DOMAIN	1312	1314	NUCLEAR	LOCALIZATION	SIGNAL	(POTENTIAL)	



```
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA DING C., POTTER E.D., QIU W., COON S.L., LEVINE M.A., GUGGINO S.E.;
RL Am. J. Physiol. 272:0-0(1997).
DR EMBL: U93851; AAC53139.1; -.
DR PFAM: PF000914; CNG_membrane; 1.
DR PROSITE: PS00888; CNMP_binding; 1.
DR PROSITE: PS00889; CNMP_BINDING; 2; 1.
DR PROSITE: PS00889; CNMP_BINDING; 2; 1.
SQ SEQUENCE 683 AA; 79167 MW; EACDI660 CRC32;

Query Match 66.0%; Score 99; DB 11; Length 683;
Best Local Similarity 50.0%; Pred. No. 0.013;
Matches 22; Conservative 7; Mismatches 1; Indels 14; Gaps 1;

QY 1 KKKKKKKK-----KKKKKKKKKKKKKKKKKK 30
   |||||:|
Db 96 KKKKKKKKADDDKESKDPKKKKKKKKKKKKKKKK 139
   |||||:|

RESULT 14
COI489 PRELIMINARY; PRT; 222 AA.
ID C01489;
AC C01489;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE CODED FOR BY C. ELEGANS CDNA YK117B5.5.
GN C13F10.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditida; Rhabditida;
OC Rhabditina; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
   elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA TIN A., WOHLDMANN P.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97006; AAC47967.1; -.
DR PFAM: PF00098; zf-CCHC; 1.
DR PFAM: PF00098; zf-CCHC; 1.
SQ SEQUENCE 222 AA; 25166 MW; F8EAC82A CRC32;

Query Match 65.3%; Score 98; DB 5; Length 222;
Best Local Similarity 63.3%; Pred. No. 0.0085;
Matches 19; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||:|
Db 127 KQKKKKKKKKKKKKKKKKKKKKKKKKKKKK 156
   |||||:|

RESULT 15
```

```
P70287 PRELIMINARY; PRT; 1587 AA.
ID P70287;
AC P70287;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE RETINOBLASTOMA BINDING PROTEIN 6 (PACT) (FRAGMENT).
GN RBBP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=TESTES;
RX MEDLINE: 97163396.
RA SIMONS A., MELAMED-BESSUDO C., WOLKOWICZ R., SPERLING J., SPERLING R.,
RA EISENBACH L., ROTTGER V.;
RT "PACT: cloning and characterization of a cellular p53 binding protein
   that interacts with Rb.";
RL Oncogene 14:145-155(1997).
DR EMBL: U28789; AAB49620.1; -.
DR MGD: MGI:894835; RBBP6.
DR PFAM: PF00097; zf-C3HC4; 1.
DR NON_TER 1
SQ SEQUENCE 1587 AA; 177272 MW; 1BE659CC CRC32;

Query Match 65.3%; Score 98; DB 11; Length 1587;
Best Local Similarity 52.2%; Pred. No. 0.027;
Matches 24; Conservative 1; Mismatches 5; Indels 16; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||:|
Db 1524 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1569
   |||||:|

Search completed: September 7, 1999, 20:34:25
Job time: 19721 sec
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:49:57 ; Search time 71.87 Seconds  
(without alignments)  
11.800 Million cell updates/sec

Title: US-09-124-280A-2  
Perfect score: 150  
Sequence: 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	72.7	511	1	NOP5_YEAST
2	108.5	72.3	441	1	PHPA_PLACH
3	108	72.0	414	1	Y694_METJA
4	107	71.3	483	1	CBF5_YEAST
5	107	71.3	504	1	SIK1_YEAST
6	106	70.7	474	1	CBF5_SCHPO
7	105	70.0	479	1	CBF5_CANAL
8	104	69.3	678	1	GARP_PLAFF
9	103	68.7	474	1	CBF5_KLULA
10	101	67.3	684	1	CNGL_MOUSE
11	101	67.3	523	1	DBP3_YEAST
12	101	67.3	1411	1	TCOF_HUMAN
13	100	66.7	686	1	CNGL_HUMAN
14	99	66.0	683	1	CNGL_RAT
15	97	64.7	690	1	CNGL_BOVIN
16	96.5	64.3	691	1	CNGL_CANFA
17	95	63.3	1178	1	MNN4_YEAST
18	95	63.3	723	1	SSRP_DROME
19	94	62.7	271	1	Y65W_YEAST
20	92.5	61.7	142	1	YMH8_YEAST
21	92	61.3	388	1	RALL_ONCVO
22	91.5	61.0	233	1	YJ08_YEAST
23	86	57.3	2231	1	SEN1_YEAST
24	86	57.3	724	1	Y061_CAEEL
25	86	57.3	344	1	YR02_YEAST
26	85	56.7	843	1	BLVR_BOVIN
27	85	56.7	645	1	CNG3_CHICK
28	85	56.7	118	1	Y093_CAEEL
29	84	56.0	1002	1	IF2P_YEAST
30	83	55.3	1213	1	YD22_DROME
31	83	55.3	167	1	YK20_YEAST
32	82.5	55.0	1240	1	VNJ1_YEAST
33	82	54.7	743	1	ABRA_PLAFC
34	82	54.7	321	1	ABRA_PLAFF
35	82	54.7	735	1	CNGL_CHICK
36	82	54.7	1005	1	YCF1_OENBE
37	81	54.0	1254	1	POLS_RRVN
38	80	53.3	487	1	CBF5_ASPTU
39	80	53.3	481	1	CBF5_EMENI
40	80	53.3	411	1	FRB3_YEAST
41	80	53.3	248	1	HI_PARAN
42	80	53.3	1220	1	IF2P_HUMAN
43	80	53.3	230	1	YD23_CAEEL

ALIGNMENTS

```
RESULT 1
NOP5_YEAST
ID NOP5_YEAST STANDARD; PRT; 511 AA.
AC Q12499;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE NUCLEOLAR PROTEIN NOP5.
GN NOP5 OR YOR310C OR O6108.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN-W303;
RX MEDLINE: 98298165.
RA WU P., BROCKENBROUGH J.S., METCALFE A.C., CHEN S., ARIS J.P.;
RT "Nop5p is a small nucleolar ribonucleoprotein component required for
pre-18S rRNA processing in yeast."
J. BIOL. CHEM. 273:16453-16463(1998).
CC -|- FUNCTION: REQUIRED FOR PRE-18S RNA PROCESSING.
CC -|- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -|- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X90565; G940841; -
DR EMBL; Z75217; E252143; -
DR EMBL; AF056070; G3170533; -
DR SGD; L0004000; NOP5.
DR NUCLEAR PROTEIN; RNA PROCESSING.
FT DOMAIN 441 511 ASP/GLU/LYS-RICH.
SQ SEQUENCE 511 AA; 56956 MW; F0A4F72C CRC32;

Query Match 72.7%; Score 109; DB 1; Length 511;
Best Local Similarity 66.7%; Pred. No. 0.00081;
Matches 20; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   ||:||||:||||:||||:||||:||||:
Db 480 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 509

RESULT 2
PHPA_PLACH
ID PHPA_PLACH STANDARD; PRT; 441 AA.
AC Q02752;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE ACIDIC PHOSPHOPROTEIN PRECURSOR (50 KD ANTIGEN).
GN PCEMAL.
OS PLASMODIUM CHABAUDI.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.
```

```

GN MJ0694.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA: EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
CC [11]
RC SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RX BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON C.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERVILAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA KLENN M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA CLOENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT *complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RT SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67516; G1591409; -.
DR TIGR: MJ0694; -.
DR HYPOTHETICAL PROTEIN.
KW DOMAIN 349 414 ASP/GLU/LYS-RICH.
FT SEQUENCE 414 AA; 47799 MW; C2CE1681 CRC32;
SQ
Query Match 72.0%; Score 108; DB 1; Length 414;
Best Local Similarity 73.3%; Pred. No. 0.00084;
Matches 22; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
DB 367 KKKKKKPKAKKKKKKKKKKKKKKKKKKK 396

RESULT 4
CBF5_YEAST STANDARD; PRT; 483 AA.
ID IC CBF5_YEAST
AC P3332;
DT 01-FEB-1994 (REL. 28, LAST CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CENTROMERE/MICROTUBULE BINDING PROTEIN CBF5 (CENTROMERE-BINDING FACTOR
DE 5) (NUCLEOLAR PROTEIN CBF5) (P64').
OS CBF5 OR YLR175W OR I9470.11.
GN SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA: FUNGI: ASCOMYCOTA: HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
CC [1]
CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE: 93330283.
RX JIANG W., MIDDLETON K., YOON H.-J., FOUQUET C., CARBON J.;
RA "An essential yeast protein, CBF5p, binds in vitro to centromeres and
RA microtubules."
RT MOL. CELL. BIOL. 13:4884-4893(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA JOHNSON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
RA KUCABA T., HALLSMORTH K., HAWKINS J., HILLIER L., JIER M.,
RA JOHNSON D., LANGSTON L., LANGSTON Y., LATREILLE P., LE T.,
RA MARDIS E., MENZIES S., MILLER N., NHAN N., PAULEY A., PELUSO D.,

```

MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
TAICH S., TREVASKIS E., VAUDIN M., VIGNATI D., WILCOX L., WILSON R.,  
WOHLMAN P., WATERSTON R.;  
SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
-|- FUNCTION: POSSIBLE MICROTUBULE BINDING PROTEIN.  
-|- SIMILARITY: BELONGS TO THE NOP5/NOP56 FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: U20237; G984964; -  
EMBL: U14913; G544506; -  
SGD: L0002982; SIKI.  
DOMAIN 443  
SEQUENCE 504 AA; 56864 MW; DF87AD52 CRC32.  
ASP/GLU/LYS-RICH.

**Qy**      1 KKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
          || ||:||:||:||:||:||:||:||  
**D6**      471 KKSKEKKEKKEKDKKEKDKKEKDKKK 500

CBF5\_SCHPO  
ID CBF5\_SCHPO STANDARD; PRT; 474 AA.  
AC O14007;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DE CENTROMERE/MICROTUBULE BINDING PROTEIN CBF5 (CENTROMERE-BINDING FACTOR  
5) (NUCLEOLAR PROTEIN CBF5).  
GN SPAC29A4.04C.

```

1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   ||| ||| ||| ||| ||| ||| ||| |||
437 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 466
Db
```

SEQUENCE FROM N.A.  
STRAIN=372;  
RA BROWN D., CHURCHER C.M., WOOD V., BARRELL B.G., RAJANDREAM M.A.;  
RC SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RL -!- FUNCTION: BINDS IN VITRO TO CENTROMERES AND MICROTUBULES. IT IS A  
CC CENTROMERE DNA-CBF3-BINDING FACTOR AND IS INVOLVED IN MITOTIC  
CC CHROMOSOME SEGREGATION. IT IS ESSENTIAL FOR CELL GROWTH (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/tutorials/licensing> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

	434	468	7 X 3 AA APPROXIMATE TANDDEM REPEATS OF K-K-E.
FT DOMAIN	434	468	
FT REPEAT	443	445	1.
FT REPEAT	450	452	2.
FT REPEAT	454	456	3.
FT REPEAT	457	459	4.

```
FT REPEAT 460 462 5.
FT REPEAT 463 465 6.
FT REPEAT 466 468 7.
SQ SEQUENCE 474 AA; 53110 MW; B544DC4A CRC32;

Query Match 70.7%; Score 106; DB 1; Length 474;
Best Local Similarity 63.3%; Pred. No. 0.0014;
Matches 19; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:||||: |||:||||: |||:||||:
Db 443 KKEKKKKKKKKKKKKKKKKKKKKKK 472

RESULT 7
CBF5_CANAL STANDARD; PRT; 479 AA.
AC O43101;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE CENTROMERE/MICROTUBULE BINDING PROTEIN CBF5 (CENTROMERE-BINDING FACTOR
DE 5) (NUCLEOLAR PROTEIN CBF5).
GN CBF5.
OS CANDIDA ALBICANS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC CANDIDACEAE; CANDIDA.
RN [1]
RP SEQUENCE FROM N.A.
RA JIANG W., CLIFFORD J., KOLTIN Y.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: BINDS IN VITRO TO CENTROMERES AND MICROTUBULES. IT IS A
CC CENTROMERE DNA-CBF3-BINDING FACTOR AND IS INVOLVED IN MITOTIC
CC CHROMOSOME SEGREGATION. IT IS ESSENTIAL FOR CELL GROWTH. MAY BE IN
CC SOME WAY ASSOCIATED WITH THE CBF3 110 KD SUBUNIT (CBF3A) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U59149; G2737890; -
KW MICROTUBULES; CENTROMERE; REPEAT; NUCLEAR PROTEIN; DNA-BINDING.
SQ SEQUENCE 479 AA; 54321 MW; 2E35229A CRC32;

Query Match 70.0%; Score 105; DB 1; Length 479;
Best Local Similarity 66.7%; Pred. No. 0.0017;
Matches 20; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:||||: |||:||||: |||:||||:
Db 432 KKEKKKKKKKKKKKKKKKKKKKKKK 461

RESULT 8
GARP_PLAFF STANDARD; PRT; 678 AA.
AC F13816;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GLUTAMIC ACID-RICH PROTEIN PRECURSOR.
GN GARP.
OS PLASMODIUM FALCIPARUM (ISOLATE FC27 / PAPUA NEW GUINEA).
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.
```

```

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE: 89040048.
RX TRIGLIA T., STAHL H.-D., CREWTER P.E., SILVA A., ANDERS R.F.,
RA KEMP D.J.;
RT "Structure of a Plasmodium falciparum gene that encodes a glutamic
RT acid-rich protein (GARP).";
RL MOL. BIOCHEM. PARASITOL. 31:199-202(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03098; G160299; -
KW REPEAT; MALARIA; ANTIGEN; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 678 GARP PROTEIN.
FT DOMAIN 120 164 15 X 3 AA TANDEM REPEATS OF K-K-X.
FT DOMAIN 372 416 9 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 417 441 5 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 576 604 POLY-GLU.
FT DOMAIN 605 653 7 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 654 663 POLY-GLU.
FT SEQUENCE 678 AA; 80551 MW; 1EE29EAE CRC32;
SQ SEQUENCE 678 AA; 80551 MW; 1EE29EAE CRC32;

Query Match 69.3%; Score 104; DB 1; Length 678;
Best Local Similarity 70.0%; Pred. No. 0.0026;
Matches 21; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:||||: |||:||||: |||:||||:
Db 119 KKKKKKKKKKKKKKKKKKKKKKKKK 148

RESULT 9
CBF5_KLUJA STANDARD; PRT; 474 AA.
AC O13473;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CENTROMERE/MICROTUBULE BINDING PROTEIN CBF5 (CENTROMERE-BINDING FACTOR
DE 5) (NUCLEOLAR PROTEIN CBF5).
DE DE 990.
GN CBF5.
OS KLUYVEROMYCES LACTIS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; KLUYVEROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JBD100;
RX MEDLINE; 98144788.
RA WINKLER A.A., BOBOK A., ZONNEVELD B.J.M., STEENSWA H.Y.,
RA HOYKAAS P.J.J.;
RT "The lysine-rich C-terminal repeats of the centromere-binding factor
RT 5 (Cbf5) of Kluyveromyces lactis are not essential for function.";
RL YEAST 14:37-48(1998).
CC -!- FUNCTION: BINDS IN VITRO TO CENTROMERES AND MICROTUBULES. IT IS A
CC CENTROMERE DNA-CBF3-BINDING FACTOR AND IS INVOLVED IN MITOTIC
CC CHROMOSOME SEGREGATION. IT IS ESSENTIAL FOR CELL GROWTH. MAY BE IN
CC SOME WAY ASSOCIATED WITH THE CBF3 110 KD SUBUNIT (CBF3A) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```





```

RX MEDLINE; 97435481.
RA RIGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";
RL YEAST 13:1077-1090(1997).
RN [3]
RP SEQUENCE OF 260-448 FROM N.A.
RX MEDLINE; 90160368.
RA CHANG T.-H., ARNAS J., ABELSON J.;
RT "Identification of five putative yeast RNA helicase genes.";
RL PROC. NATL. ACAD. SCI. U.S.A. 87:1571-1575(1990).
CC -!- FUNCTION: INVOLVED IN RIBOSOME BIOGENESIS AND MORE SPECIFICALLY IN
CC A STEP LEADING TO THE PRODUCTION OF 5.8S AND 25S RRNA AT THE LEVEL
CC OF THE PROCESSING OF THE 27S RRNA PRECURSOR.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M80437; G172582; -.
DR EMBL; 272600; E243966; -.
DR PIR; B34848; B34848.
DR SGD; L0000494; DBP3.
DR PFAM; PF00270; DEAD; 1.
DR PFAM; PF00271; helicase.C; 1.
KW HELICASE; ATP-BINDING; RNA-BINDING; NUCLEAR PROTEIN; RNA PROCESSING.
FT NP_BIND 156 163
FT SITE 262 265
FT CONFLICT 441 441 E -> A (IN REF. 3).
FT CONFLICT 444 444 V -> I (IN REF. 3).
SQ SEQUENCE 523 AA; 58826 MW; D6005B53 CRC32;

Query Match 67.3%; Score 101; DB 1; Length 523;
Best Local Similarity 70.0%; Pred. No. 0.0039;
Matches 21; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30
Db 21 KKKKKKKKKKKKKKKKKKKKKKKKKKK 50

RESULT 12
TCOF HUMAN
ID TCOF HUMAN STANDARD; PRT; 1411 AA.
AC Q13428; Q99408;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TREACLE PROTEIN (TREACHER COLLINS SYNDROME PROTEIN).
GN TCOF1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96154183.
RA DIXON J., EDWARDS S.J., GLADWIN A.J., DIXON M.J., LOFTUS S.K.,
RA BONNER C.A., KOPRIVNIKAR K., WASMUTH J.J.;
RT "Positional cloning of a gene involved in the pathogenesis of
RT Treacher Collins syndrome. The Treacher Collins Syndrome
RT Collaborative Group.";
RL NAT. GENET. 12:130-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97250498.

```

```

RA WISE C.A., CHIANG L.C., PAZNEKAS W.A., SHARMA M., MUSY M.M.,
RA ASHLEY J.A., LOVETT M., JABS E.W.;
RT "TCOF1 gene encodes a putative nucleolar phosphoprotein that exhibits
RT mutations in Treacher Collins Syndrome throughout its coding
RT region.";
RL PROC. NATL. ACAD. SCI. U.S.A. 94:3110-3115(1997).
RN [3]
RP VARIANTS L-439; V-810; V-1313 & G-1355, AND VARIANT TCS R-53.
RX MEDLINE; 97195537.
RA EDWARDS S.J., GLADWIN A.J., DIXON M.J.;
RT "The mutational spectrum in Treacher Collins syndrome reveals a
RT predominance of mutations that create a premature-termination
RT codon.";
RL AM. J. HUM. GENET. 60:515-524(1997).
CC -!- DISEASE: DEFECTS IN TCOF1 ARE THE CAUSE OF TREACHER COLLINS
CC SYNDROME (TCS). TCS IS A AUTOSOMAL DOMINANT DISORDER OF
CC CRANIOFACIAL DEVELOPMENT THAT OCCURS WITH AN INCIDENCE OF 1/50,000
CC LIVE BIRTHS. THE CLINICAL FEATURES OF TCS ARE BILATERALLY
CC SYMMETRICAL AND INCLUDE: (1) ABNORMALITIES OF THE EXTERNAL EARS,
CC ATRESIA OF THE EXTERNAL EAR CANALS, AND MALFORMATION OF THE MIDDLE
CC EAR OSSICLES, WHICH MAY RESULT IN CONDUCTIVE HEARING LOSS; (2)
CC LATERAL DOWNSLOPING OF PALPEBRAL FISSURES, FREQUENTLY WITH
CC COLOBOMAS OF THE LOWER EYELIDS; (3) HYPOPLASIA OF THE MANDIBLE AND
CC ZYGOMATIC COMPLEX; (4) CLEFT PALATE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U40847; G1736917; -.
DR EMBL; U76366; G1685049; -.
DR EMBL; U84664; G1854952; -.
DR EMBL; U84640; G1854952; JOINED.
DR EMBL; U84641; G1854952; JOINED.
DR EMBL; U84642; G1854952; JOINED.
DR EMBL; U84643; G1854952; JOINED.
DR EMBL; U84644; G1854952; JOINED.
DR EMBL; U84645; G1854952; JOINED.
DR EMBL; U84646; G1854952; JOINED.
DR EMBL; U84647; G1854952; JOINED.
DR EMBL; U84648; G1854952; JOINED.
DR EMBL; U84649; G1854952; JOINED.
DR EMBL; U84650; G1854952; JOINED.
DR EMBL; U84651; G1854952; JOINED.
DR EMBL; U84652; G1854952; JOINED.
DR EMBL; U84653; G1854952; JOINED.
DR EMBL; U84654; G1854952; JOINED.
DR EMBL; U84655; G1854952; JOINED.
DR EMBL; U84656; G1854952; JOINED.
DR EMBL; U84657; G1854952; JOINED.
DR EMBL; U84658; G1854952; JOINED.
DR EMBL; U84659; G1854952; JOINED.
DR EMBL; U84660; G1854952; JOINED.
DR EMBL; U84661; G1854952; JOINED.
DR EMBL; U84662; G1854952; JOINED.
DR EMBL; U84663; G1854952; JOINED.
DR MIM; 154500; -.
KW DISEASE MUTATION; POLYMORPHISM.
RN [1]
RP DOMAIN 89 97
RX MEDLINE; 96154183.
RA DIXON J., EDWARDS S.J., GLADWIN A.J., DIXON M.J., LOFTUS S.K.,
RA BONNER C.A., KOPRIVNIKAR K., WASMUTH J.J.;
RT "Positional cloning of a gene involved in the pathogenesis of
RT Treacher Collins syndrome. The Treacher Collins Syndrome
RT Collaborative Group.";
RL NAT. GENET. 12:130-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97250498.

```

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announcement/send_an_email_to_license@isb-sib.ch)).  
 -----

# CC -!- FUNCTION: VISUAL SIGNAL TRANSDUCTION IS MEDIATED BY A G-PROTEIN



Job time: 1938 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:12 ; Search time 116.8 Seconds  
(without alignments)  
10.291 Million cell updates/sec

Title: US-09-124-280A-2  
Perfect score: 150  
Sequence: 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	74.0	215	2 I52523	nucleoporin p62 ho
2	109	72.7	511	2 S58322	nucleolar protein
3	108.5	72.3	441	2 A48455	acidic phosphoprot
4	108	72.0	414	2 F64386	hypothetical prote
5	108	72.0	812	2 S43604	R07ES.1 protein (c
6	107	71.3	504	2 S48550	hypothetical prote
7	107	71.3	483	2 S41853	centromere/microtu
8	104	69.3	678	2 A54514	glutamic acid-rich
9	103	68.7	508	2 E71620	hypothetical prote
10	101.5	67.7	1320	2 JC5630	TCOF1 protein - mo
11	101	67.3	523	2 S30805	probable RNA helic
12	101	67.3	688	2 B42161	cGMP-gated cation
13	100	66.7	686	2 A44842	cGMP-gated ion cha
14	100	66.7	690	2 A42161	cGMP-gated cation
15	97	64.7	690	2 S07103	cAMP-gated ion cha
16	96.5	64.3	691	2 JC6509	rod cyclic nucleot
17	95	63.3	723	2 S33688	hypothetical prote
18	95	63.3	723	2 A48217	single-strand DNA/
19	95	63.3	1178	2 S78475	mannosylphosphoryl
20	95	63.3	483	2 E71619	hypothetical prote
21	94	62.7	271	2 S64615	hypothetical prote
22	94	62.7	754	2 JC5314	CDC28/cdc2-like ki
23	92.5	61.7	142	2 S54481	hypothetical prote
24	92	61.3	336	2 A32507	41k larval antigen
25	91.5	61.0	233	2 S55165	hypothetical prote
26	91	60.7	516	2 D64410	replication factor
27	89	59.3	751	2 S65459	DNA topoisomerase
28	89	59.3	189	2 T02570	hypothetical prote
29	87	58.0	405	2 S42875	dihydrolipoamide S
30	86	57.3	2231	2 S53416	SEN1 protein - yea
31	86	57.3	344	2 S45912	YRO2 protein - yea
32	86	57.3	695	2 S40923	hypothetical prote
33	85	56.7	645	2 I50680	alpha subunit of r
34	85	56.7	972	2 F71608	hypothetical prote
35	85	56.7	729	2 A45716	leukemia virus cel
36	84.5	56.3	1095	2 T00329	hypothetical prote
37	84	56.0	605	2 T02350	hypothetical prote
38	84	56.0	683	2 T00872	probable receptor
39	84	56.0	1002	2 S70292	FUN12 protein - ye

probable membrane  
hypothetical prote  
hypothetical prote  
TATA-binding prote  
hypothetical prote  
alpha subunit of c

ALIGNMENTS

RESULT 1  
I52523  
nucleoporin p62 homolog - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 27-Feb-1997  
C:Accession: I52523  
R:Wang, Z.Q.; Akmal, K.M.; Kim, K.H.  
Biol. Reprod. 51, 1022-1030, 1994  
A:Title: An unusual nucleoporin-related messenger ribonucleic acid is present in the  
A:Reference number: I52523; MUID:95151924  
A:Accession: I52523  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-215 <RES>  
A:Cross-references: GB:S75997; NID:913245; PID:g913246  
A:Experimental source: testis

Query Match 74.0% Score 111; DB 2; Length 215;  
Best Local Similarity 95.7%; Pred. NO. 0.00043;  
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKK 23  
:|||||:|||||:|||||:|||||:|||||  
Db 34 EKKKKKKKKKKKKKKKKKKKKKKKK 56

RESULT 2  
S58322  
nucleolar protein NOP5 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein O6108; protein YOR310c  
C:Species: Saccharomyces cerevisiae  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 24-Sep-1998  
C:Accession: S58322; S67216; S71990  
R:Pearson, B.M.; Hernando, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: S58318  
A:Accession: S58322  
A:Molecule type: DNA  
A:Residues: 1-511 <PEA>  
A:Cross-references: EMBL:X90565; NID:g940836; PID:g940841  
R:Pearson, B.M.; Hernando, Y.; Kalogeropoulos, A.; Schweizer, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67213  
A:Accession: S67216  
A:Molecule type: DNA  
A:Residues: 1-511 <PEW>  
A:Cross-references: EMBL:Z75217; NID:g1420680; PID:e252143; PID:g1420682; MIPS:YOR310  
A:Experimental source: strain S288c  
R:Pearson, B.M.; Hernando, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.  
Yeast 12, 1021-1031, 1996  
A:Title: Sequencing of a 35.71 kb DNA segment on the right arm of yeast chromosome XV  
A:Reference number: S71986  
A:Accession: S71990  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-511 <PEF>  
A:Cross-references: EMBL:X90565; NID:g940836; PID:g940841  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995  
C:Genetics:  
A:Gene: SGD:NOP5  
A:Cross-references: SGD:S0005837; MIPS:YOR310c

```
Qy      1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
        |||||: | |||: | |||: | |||: | |||: | |||: | |||: | |||
```

A;Mole

S41853  
centromere/microtubule-binding protein CBF5 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein L9470.11; protein YLR175W  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Mar-1992 #sequence\_revision 14-Sep-1994 #text\_change 06-Feb-1998  
C:Accession: S41853; S45443; S51420  
R:Jiang, W.; Middleton, K.; Yoon, H.J.; Fouquet, C.; Carbon, J.  
Mol. Cell. Biol. 13, 4884-4893, 1993  
A:Title: An essential yeast protein, CBF5p, binds in vitro to centromeres and  
A:Reference number: S41853; M0ID:93330283  
A:Accession: S41853  
A:Molecule type: DNA



```
A:Residues: 1-483 <JTA>
A:Cross-references: EMBL:L12351; NID:g311096; PID:g311097
A:Accession: S45443
A:Molecule type: protein
A:Residues: 36-45;142-149;168-175;385-401 <J12>
R:Wohldmann, P.
Submitted to the EMBL Data Library, November 1994
A:Description: The sequence of S. cerevisiae cosmid 9470.
A:Reference number: S51414
A:Accession: S51420
A:Molecule type: DNA
A:Residues: 1-483 <WOH>
A:Cross-references: EMBL:U17246; NID:g577192; PID:g577203; MIPS:YLR175w
C:Genetics:
A:Gene: SGD:CBF5
A:Cross-references: SGD:S0004165; MIPS:YLR175w
A:Map position: 12R
F:434-463/Region: Lys-Lys-Glu/Asp repeats

Query Match      71.3%; Score 107; DB 2; Length 483;
Best Local Similarity 66.7%; Pred. No. 0.0016;
Matches 20; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:|||:|||:|||:|||:|||:|||:|||
Db 437 KKEKKKKKKKKKKKKKKKKKKKKKKKKR 466

RESULT 8
A54514
glutamic acid-rich protein precursor - Plasmodium falciparum
N:Alternate names: GARP
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 09-Sep-1997
C:Accession: A54514
R:Triglia, T.; Stahl, H.D.; Crewther, P.E.; Silva, A.; Anders, R.F.; Kemp, D.J.
Mol. Biochem. Parasitol. 31, 199-202, 1988
A:Title: Structure of a Plasmodium falciparum gene that encodes a glutamic acid-rich pro
A:Reference number: A54514
A:Accession: A54514
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-678 <TRI>
A:Cross-references: GB:J03998; NID:g160298; PID:g160299
C:Genetics:
A:Introns: 25/3
C:Keywords: tandem repeat

Query Match      69.3%; Score 104; DB 2; Length 678;
Best Local Similarity 70.0%; Pred. No. 0.0035;
Matches 21; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:|||:|||:|||:|||:|||:|||:|||
Db 119 KKKKKKKKKKKKKKKKKKKKKKKKKKK 148

RESULT 9
E71620
hypothetical protein PF80225c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: E71620
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: E71620
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-508 <GAR>
A:Cross-references: GB:AE001381; GB:AE001362; NID:g3845124; PID:g3845128; TIGR:PF80222
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PF80225c

Query Match      68.7%; Score 103; DB 2; Length 508;
Best Local Similarity 63.3%; Pred. No. 0.0035;
Matches 19; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||:|||:|||:|||:|||:|||:|||:|||
Db 390 KKEKKKKKKKKKKKKKKKKKKKKKKKEK 419

RESULT 10
JC5630
TCOF1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 17-Mar-1999
C:Accession: JC5630
R:Paznekas, W.A.; Zhang, N.; Gridley, T.; Jabbs, E.W.
Biochem. Biophys. Res. Commun. 238, 1-6, 1997
A:Title: Mouse TCOF1 is expressed widely, has motifs conserved in nucleolar phosphop
A:Reference number: JC5630; MUID:97445113
A:Accession: JC5630
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1320 <PAZ>
A:Cross-references: DDBJ:AF001794; NID:g2109458; PID:g2109459
C:Comment: This protein is a nucleolar phosphoprotein with 82 potential phosphorylati
ing.

Query Match      67.7%; Score 101.5; DB 2; Length 1320;
Best Local Similarity 58.5%; Pred. No. 0.0086;
Matches 24; Conservative 2; Mismatches 4; Indels 11; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||:|||:|||:|||:|||:|||:|||
Db 1274 KKKSLAKKKKKKKKKKKKKKKKKKKKK 1314

RESULT 11
S30805
probable RNA helicase CA3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G3210; protein YGL078c
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 13-Sep-1998
C:Accession: S30805; B34848; S64085
R:Johnston, M.; Nogae, I.
Submitted to the EMBL Data Library, December 1991
A:Reference number: S30805
A:Accession: S30805
A:Molecule type: DNA
A:Residues: 1-523 <JOH>
A:Cross-references: EMBL:M80437; NID:g172581; PID:g172582
R:Chang, T.H.; Arenas, J.; Abelson, J.
Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990
A:Title: Identification of five putative yeast RNA helicase genes.
A:Reference number: A34848; MUID:90160368
A:Accession: B34848
A:Molecule type: DNA
A:Residues: 260-440, A', 442-443, I', 445-448 <CHA>
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64085
A:Molecule type: DNA
A:Residues: 1-523 <RIE>
A:Cross-references: EMBL:Z72600; NID:g1322594; PID:e243966; PID:g1322595; MIPS:YGL078
A:Experimental source: strain S288c
```

## C:Genetics:

A:Gene: SGD:DBP3  
A:Cross-references: SGD:S0003046; MIPS:YGL078c  
A:Map position: 7L  
C:Superfamily: ATP-dependent RNA helicase DBP1; DEAD/H box helicase homology  
C:Keywords: ATP; P-loop  
F:156-452/Domain: DEAD/H box helicase homology <DEAD>  
F:156-163/Region: nucleotide-binding motif A (P-loop)  
F:258-263/Region: nucleotide-binding motif B  
F:262-265/Region: DEAD motif

Query Match 67.3%; Score 101; DB 2; Length 523;  
Best Local Similarity 70.0%; Pred. No. 0.0052;  
Matches 21; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 21 KKKKKKKKKKKKKKKKKKKKKKKKKKK 50

## RESULT 12

CGMP-gated cation channel, rod photoreceptor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Feb-1997  
C:Accession: B42161  
R:Pittler, S.J.; Lee, A.K.; Altherr, M.R.; Howard, T.A.; Seldin, M.F.; Hurwitz, R.L.; We  
J. Biol. Chem. 267, 6257-6262, 1992  
A:Title: Primary structure and chromosomal localization of human and mouse rod photorece  
A:Reference number: A42161; MUID:92210603  
A:Accession: B42161  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-688 <PIT>  
A:Cross-references: GB:M84742  
A:Note: authors translated the codon TAT for residue 544 as Thr  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-  
C:Keywords: CGMP binding  
F:475-599/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA2>

Query Match 67.3%; Score 101; DB 2; Length 688;  
Best Local Similarity 52.4%; Pred. No. 0.0063;  
Matches 22; Conservative 7; Mismatches 1; Indels 12; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 98 KKKKKKKKKKKKKKKKKKKKKKKKKKK 139

## RESULT 13

CGMP-gated ion channel protein - human  
N:Alternate names: rod photoreceptor CGMP-gated channel  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Mar-1998  
C:Accession: A44842  
R:Dhallan, R.S.; Macke, J.P.; Eddy, R.L.; Shows, T.B.; Reed, R.R.; Yau, K.W.; Nathans, J  
J. Neurosci. 12, 3248-3256, 1992  
A:Title: Human rod photoreceptor CGMP-gated channel: amino acid sequence, gene structure  
A:Reference number: A44842; MUID:92356211  
A:Accession: A44842  
A:Molecule type: mRNA  
A:Residues: 1-686 <DHA>  
A:Cross-references: GB:S42457; NID:q252853; PID:q252854  
A:Experimental source: retina  
A:Note: sequence extracted from NCBI backbone (NCBIN:110250, NCBIP:110251)  
C:Note: intron positions were determined from genomic sequence  
C:Genetics:  
A:Gene: GDB:CNCG1; CNCG  
A:Cross-references: GDB:127557; OMIM:123825  
A:Map position: 4p14-q13

A:Introns: 36/2; 75/2; 95/2; 110/2; 146/2; 182/2; 218/2  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleot  
C:Keywords: CGMP binding; eye; ion channel; ion transport; retina; transmembrane prot  
F:475-599/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 66.7%; Score 100; DB 2; Length 686;  
Best Local Similarity 52.3%; Pred. No. 0.0076;  
Matches 23; Conservative 4; Mismatches 3; Indels 14; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 101 KKKKKKKKKKKKKKKKKKKKKKKKKKK 144

## RESULT 14

CGMP-gated cation channel, rod photoreceptor - human  
C:Species: Homo sapiens (man)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 08-Sep-1997  
C:Accession: A42161  
R:Pittler, S.J.; Lee, A.K.; Altherr, M.R.; Howard, T.A.; Seldin, M.F.; Hurwitz, R.L.;  
J. Biol. Chem. 267, 6257-6262, 1992  
A:Title: Primary structure and chromosomal localization of human and mouse rod photor  
A:Reference number: A42161; MUID:92210603  
A:Accession: A42161  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-690 <PIT>  
A:Cross-references: GB:M84741; NID:gl80461; PID:gl80462  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleot  
C:Keywords: CGMP binding; transmembrane protein  
F:479-603/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 66.7%; Score 100; DB 2; Length 690;  
Best Local Similarity 52.3%; Pred. No. 0.0076;  
Matches 23; Conservative 4; Mismatches 3; Indels 14; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 105 KKKKKKKKKKKKKKKKKKKKKKKKKKK 148

## RESULT 15

CGMP-gated ion channel protein - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 29-Jan-1999  
C:Accession: S07103  
R:Kaupp, U.B.; Niidome, T.; Tanabe, T.; Terada, S.; Boenigk, W.; Stuehmer, W.; Cook,  
Nature 342, 762-766, 1989  
A:Title: Primary structure and functional expression from complementary DNA of the ro  
A:Reference number: S07103; MUID:90098076  
A:Accession: S07103  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-690 <KAU>  
A:Cross-references: GB:X51604; NID:g203; PID:g204  
A:Note: part of this sequence was confirmed by protein sequencing  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleot  
C:Keywords: CGMP binding; eye; ion channel; ion transport; retina; transmembrane prot  
F:477-601/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 64.7%; Score 97; DB 2; Length 690;  
Best Local Similarity 52.4%; Pred. No. 0.014;  
Matches 22; Conservative 5; Mismatches 3; Indels 12; Gaps 1;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
||| ||| ||| ||| ||| ||| ||| ||| |||  
DB 100 KKKKKKKKKKKKKKKKKKKKKKKKKKK 141

Search completed: September 7, 1999, 23:06:13  
Job time: 2461 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:36:59 ; Search time 147.16 Seconds  
(without alignments)  
4.829 Million cell updates/sec

Title: US-09-124-280A-2  
Perfect score: 150  
Sequence: 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150	100.0	30	1 W21590	Antibiotic potenti
2	150	100.0	434	1 W21591	Antibiotic potenti
3	150	100.0	45	1 W24450	Nucleic acid (NA)
4	150	100.0	32	1 W38843	Delivery peptide u
5	150	100.0	33	1 W38844	Delivery peptide u
6	150	100.0	34	1 W38845	Delivery peptide u
7	150	100.0	35	1 W38846	Delivery peptide u
8	150	100.0	32	1 W38806	Delivery peptide u
9	150	100.0	33	1 W38807	Delivery peptide u
10	150	100.0	34	1 W38808	Delivery peptide u
11	150	100.0	36	1 W38847	Delivery peptide u
12	150	100.0	37	1 W38848	Delivery peptide u
13	150	100.0	38	1 W38849	Delivery peptide u
14	150	100.0	39	1 W38850	Delivery peptide u
15	150	100.0	40	1 W38851	Delivery peptide u
16	150	100.0	41	1 W38852	Delivery peptide u
17	150	100.0	33	1 W38887	Delivery peptide u
18	150	100.0	42	1 W38853	Delivery peptide u
19	150	100.0	34	1 W38888	Delivery peptide u
20	150	100.0	43	1 W38854	Delivery peptide u
21	150	100.0	35	1 W38889	Delivery peptide u
22	150	100.0	36	1 W38890	Delivery peptide u
23	150	100.0	37	1 W38891	Delivery peptide u
24	150	100.0	38	1 W38892	Delivery peptide u
25	150	100.0	39	1 W38893	Delivery peptide u
26	150	100.0	40	1 W38894	Delivery peptide u
27	150	100.0	41	1 W38895	Delivery peptide u
28	150	100.0	42	1 W38896	Delivery peptide u
29	150	100.0	43	1 W38897	Delivery peptide u
30	150	100.0	44	1 W38898	Delivery peptide u
31	150	100.0	35	1 W38899	Delivery peptide u
32	150	100.0	36	1 W38910	Delivery peptide u
33	150	100.0	37	1 W38811	Delivery peptide u
34	150	100.0	38	1 W38812	Delivery peptide u
35	150	100.0	39	1 W38813	Delivery peptide u
36	150	100.0	40	1 W38814	Delivery peptide u
37	150	100.0	41	1 W38815	Delivery peptide u
38	150	100.0	42	1 W38816	Delivery peptide u
39	150	100.0	43	1 W38817	Delivery peptide u
40	150	100.0	40	1 W65939	Polylysine peptide
41	145	96.7	31	1 W38842	Delivery peptide u
42	145	96.7	31	1 W38805	Delivery peptide u
43	145	96.7	32	1 W38886	Delivery peptide u

ALIGNMENTS

RESULT 1

W21590 ID W21590 standard; peptide; 30 AA.  
AC W21590;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #2.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI: 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Disclosure; Page 7; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 30 AA;

Query Match 100.0%; Score 150; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 7.9e-10;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 2

W21591 ID W21591 standard; peptide; 434 AA.

AC W21591;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #3.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

PN WO9638163-A1.

PD 05-DEC-1996.

PF 23-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI: 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Disclosure; Page 7; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to

CC potentiate the activity of an antibiotic when they are co-administered

CC with the antibiotic. Compositions containing these peptides are used

CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction

CC in the dose of antibiotic required by 10-90% of the normal dose for

CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 434 AA;

Query Match 100.0%; Score 150; DB 1; Length 434;  
Best Local Similarity 100.0%; Pred. No. 5.2e-09;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 3  
ID W24450 standard; peptide; 45 AA.  
AC W24450;  
DT 30-SEP-1997 (first entry)  
DE Nucleic acid (NA) binding peptide used in NA delivery to cells.  
KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;  
KW JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.  
OS Synthetic.  
PN WO9640958-A1.  
PD 19-DEC-1996.  
PF 23-APR-1996; U05679.  
PR 07-JUN-1995; US-484777.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Smith LC, Sparrow JT, Woo SL;  
DR WPI; 97-052345/05.  
PT Nucleic acid transporter useful in gene therapy - contains binding complex associated with surface and nuclear ligands and lysis agent  
PS Disclosure; Page 49; 125pp; English.  
CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both condensing and stabilising a NA. The peptides can be conjugated to a lytic peptide to form a nucleic acid transporter system. The lysis agent forms an alpha-helical structure. The transporter system is used to deliver nucleic acid to a cell and for treating humans by gene therapy. By taking advantage of the characteristics of both the lysis agents and the binding molecules, delivery of the nucleic acid is enhanced. Specific lysis agents are capable of releasing the nucleic acid into the cellular interior from the endosome. Release is efficient without endosomal/lysosomal degradation. Once released the binding complexes help target the nucleic acid to the nucleus.  
SQ Sequence 45 AA;

Query Match 100.0%; Score 150; DB 1; Length 45;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
| | | | | | | | | | | | | | | | | | | | | |  
Db 4 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 33

RESULT 4  
ID W38843 standard; peptide; 32 AA.  
AC W38843;  
DT 30-MAR-1998 (first entry)  
DE Delivery peptide; peptide-macromolecule complex; macromolecule delivery;  
KW Delivery peptide; peptide-macromolecule complex; macromolecule delivery;  
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc\_difference 31 /note= "any amino acid"  
PN WO9725070-A2.  
PD 17-JUL-1997.  
PF 02-JAN-1997; U00454.  
PR 08-JAN-1996; US-584043.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.

PI Hauer J, Mims MP, Smith LC, Sparrow JT;  
DR WPI; 97-372622/34.  
PT New lipophilic peptide-macromolecule complexes - used for the delivery of macromolecules to cells, particularly for gene therapy  
PS Claim 6; Page 71; 106pp; English.  
CC This sequence represents a delivery peptide that can be used in the peptide-macromolecule complex of the invention. The peptide-macromolecule complex of the invention is for delivering a macromolecule into a cell, and comprises a non-exchangeable lipophilic peptide (LP) comprising a delivery peptide associated with a lipid moiety, where the delivery peptide portion of the LP is complexed to the macromolecule. The complexes can be used for the delivery of macromolecules such as nucleic acids, proteins, oligonucleotides, lipids or carbohydrates. They can be used to treat diseases by enhancing delivery of specific nucleic acid to the appropriate targeted cells. They can also be used to create transformed cells as well as transgenic animals for assessing human disease in an animal model. They can also be used for livestock agricultural purposes. The complex is capable of transporting the macromolecule in a stable and condensed state and releasing the molecule into the cellular interior. The complex can bind with a cell surface receptor, lyse an endosome and target the nucleus of the cell.  
SQ Sequence 32 AA;

Query Match 100.0%; Score 150; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 8.2e-10;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 5  
ID W38844 standard; peptide; 33 AA.  
AC W38844;  
DT 30-MAR-1998 (first entry)  
DE Delivery peptide used in peptide macromolecule complex.  
KW Delivery peptide; peptide-macromolecule complex; macromolecule delivery;  
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc\_difference 32 /note= "any amino acid"  
PN WO9725070-A2.  
PD 17-JUL-1997.  
PF 02-JAN-1997; U00454.  
PR 08-JAN-1996; US-584043.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Hauer J, Mims MP, Smith LC, Sparrow JT;  
DR WPI; 97-372622/34.  
PT New lipophilic peptide-macromolecule complexes - used for the delivery of macromolecules to cells, particularly for gene therapy  
PS Claim 6; Page 71; 106pp; English.  
CC This sequence represents a delivery peptide that can be used in the peptide-macromolecule complex of the invention. The peptide-macromolecule complex of the invention is for delivering a macromolecule into a cell, and comprises a non-exchangeable lipophilic peptide (LP) comprising a delivery peptide associated with a lipid moiety, where the delivery peptide portion of the LP is complexed to the macromolecule. The complexes can be used for the delivery of macromolecules such as nucleic acids, proteins, oligonucleotides, lipids or carbohydrates. They can be used to treat diseases by enhancing delivery of specific nucleic acid to the appropriate targeted cells. They can also be used to create transformed cells as well as transgenic animals for assessing human disease in an animal model. They can also be used for livestock agricultural purposes. The complex is capable of transporting the macromolecule in a stable and condensed state and releasing the molecule into the cellular interior. The complex can bind with a cell surface receptor, lyse an endosome and target the nucleus of the cell.  
SQ Sequence 33 AA;

```

Query Match      100.0%; Score 150; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 6
W38845
ID W38845 standard; peptide; 34 AA.
AC W38845;
DT 30-MAR-1998 (first entry)
DE Delivery peptide used in peptide macromolecule complex.
KW Delivery peptide; peptide-macromolecule complex; macromolecule delivery;
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 33 /note= "any amino acid"
PN WO9725070-A2.
PD 17-JUL-1997.
PF 02-JAN-1997; U00454.
PR 08-JAN-1996; US-584043.
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Hauer J, Mims MP, Smith LC, Sparrow JT;
DR WPI: 97-372622/34.
PT New lipophilic peptide-macromolecule complexes - used for the
PT delivery of macromolecules to cells, particularly for gene therapy
PS Claim 6; Page 72; 106pp; English.
CC This sequence represents a delivery peptide that can be used in the
CC peptide-macromolecule complex of the invention. The peptide-macromolecule
CC complex of the invention is for delivering a macromolecule into a cell,
CC and comprises a non-exchangeable lipophilic peptide (LP) comprising a
CC delivery peptide associated with a lipid moiety, where the delivery
CC peptide portion of the LP is complexed to the macromolecule. The
CC complexes can be used for the delivery of macromolecules such as nucleic
CC acids, proteins, oligonucleotides, lipids or carbohydrates. They can be
CC used to treat diseases by enhancing delivery of specific nucleic acid to
CC the appropriate targeted cells. They can also be used to create
CC transformed cells as well as transgenic animals for assessing human
CC disease in an animal model. They can also be used for livestock
CC agricultural purposes. The complex is capable of transporting the
CC macromolecule in a stable and condensed state and releasing the molecule
CC into the cellular interior. The complex can bind with a cell surface
CC receptor, lyse an endosome and target the nucleus of the cell.
SQ Sequence 35 AA;

Query Match      100.0%; Score 150; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 8
W38806
ID W38806 standard; peptide; 32 AA.
AC W38806;
DT 30-MAR-1998 (first entry)
DE Delivery peptide used in peptide macromolecule complex.
KW Delivery peptide; peptide-macromolecule complex; macromolecule delivery;
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.
OS Synthetic.
PN WO9725070-A2.
PD 17-JUL-1997.
PF 02-JAN-1997; U00454.
PR 08-JAN-1996; US-584043.
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Hauer J, Mims MP, Smith LC, Sparrow JT;
DR WPI: 97-372622/34.
PT New lipophilic peptide-macromolecule complexes - used for the
PT delivery of macromolecules to cells, particularly for gene therapy
PS Claim 6; Page 89; 106pp; English.
CC This sequence represents a delivery peptide that can be used in the
CC peptide-macromolecule complex of the invention. The peptide-macromolecule
CC complex of the invention is for delivering a macromolecule into a cell,
CC and comprises a non-exchangeable lipophilic peptide (LP) comprising a
CC delivery peptide associated with a lipid moiety, where the delivery
CC peptide portion of the LP is complexed to the macromolecule. The
CC complexes can be used for the delivery of macromolecules such as nucleic
CC acids, proteins, oligonucleotides, lipids or carbohydrates. They can be
CC used to treat diseases by enhancing delivery of specific nucleic acid to
CC the appropriate targeted cells. They can also be used to create
CC transformed cells as well as transgenic animals for assessing human
CC disease in an animal model. They can also be used for livestock
CC agricultural purposes. The complex is capable of transporting the
CC macromolecule in a stable and condensed state and releasing the molecule
CC into the cellular interior. The complex can bind with a cell surface
CC receptor, lyse an endosome and target the nucleus of the cell.
SQ Sequence 32 AA;

Query Match      100.0%; Score 150; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||||
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 7
W38846
ID W38846 standard; peptide; 35 AA.
AC W38846;
DT 30-MAR-1998 (first entry)
DE Delivery peptide used in peptide macromolecule complex.
KW Delivery peptide; peptide-macromolecule complex; macromolecule delivery;
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 34 /note= "any amino acid"
PN WO9725070-A2.

```





PF 02-JAN-1997; U00454.  
PR 08-JAN-1996; US-584043.  
PI (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Hauer J, Mims MP, Smith LC, Sparrow JT;  
PI WPI: 97-372622/34.  
PT New lipophilic peptide-macromolecule complexes - used for the  
PT delivery of macromolecules to cells, particularly for gene therapy  
PT Claim 6; Page 74; 106pp; English.  
PS This sequence represents a delivery peptide that can be used in the  
CC peptide-macromolecule complex of the invention. The peptide-macromolecule  
CC complex of the invention is for delivering a macromolecule into a cell,  
CC and comprises a non-exchangeable lipophilic peptide (LP) comprising a  
CC delivery peptide associated with a lipid moiety, where the delivery  
CC peptide portion of the LP is complexed to the macromolecule. The  
CC complexes can be used for the delivery of macromolecules such as nucleic  
CC acids, proteins, oligonucleotides, lipids or carbohydrates. They can be  
CC used to treat diseases by enhancing delivery of specific nucleic acid to  
CC the appropriate targeted cells. They can also be used to create  
CC transformed cells as well as transgenic animals for assessing human  
CC disease in an animal model. They can also be used for livestock  
CC agricultural purposes. The complex is capable of transporting the  
CC macromolecule in a stable and condensed state and releasing the molecule  
CC into the cellular interior. The complex can bind with a cell surface  
CC receptor, lyse an endosome and target the nucleus of the cell.  
SQ Sequence 38 AA;

Query Match 100.0%; Score 150; DB 1; Length 38;  
Best Local Similarity 100.0%; Pred. NO. 9.3e-10;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII  
DB 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 14  
W38850  
ID W38850 standard; peptide; 39 AA.  
AC W38850;  
AC 30-MAR-1998 (first entry)  
DE Delivery peptide used in peptide macromolecule complex.  
DT Delivery peptide; peptide-macromolecule complex; macromolecule delivery;  
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.  
KW Synthetic.  
OS  
FH Key Location/Qualifiers  
FT Misc\_difference 38  
FT /note= "any amino acid"  
FN W09725070-A2.  
PT 17-JUL-1997.  
PT 02-JAN-1997; U00454.  
PR 08-JAN-1996; US-584043.  
PI (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Hauer J, Mims MP, Smith LC, Sparrow JT;  
PI WPI: 97-372622/34.  
PT New lipophilic peptide-macromolecule complexes - used for the  
PT delivery of macromolecules to cells, particularly for gene therapy  
PT Claim 6; Page 74; 106pp; English.  
PS This sequence represents a delivery peptide that can be used in the  
CC peptide-macromolecule complex of the invention. The peptide-macromolecule  
CC complex of the invention is for delivering a macromolecule into a cell,  
CC and comprises a non-exchangeable lipophilic peptide (LP) comprising a  
CC delivery peptide associated with a lipid moiety, where the delivery  
CC peptide portion of the LP is complexed to the macromolecule. The  
CC complexes can be used for the delivery of macromolecules such as nucleic  
CC acids, proteins, oligonucleotides, lipids or carbohydrates. They can be  
CC used to treat diseases by enhancing delivery of specific nucleic acid to  
CC the appropriate targeted cells. They can also be used to create  
CC transformed cells as well as transgenic animals for assessing human  
CC disease in an animal model. They can also be used for livestock  
CC agricultural purposes. The complex is capable of transporting the  
CC macromolecule in a stable and condensed state and releasing the molecule  
CC into the cellular interior. The complex can bind with a cell surface

CC receptor, lyse an endosome and target the nucleus of the cell.  
SQ Sequence 39 AA;

Query Match 100.0%; Score 150; DB 1; Length 39;  
Best Local Similarity 100.0%; Pred. No. 9.5e-10;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
| | | | | | | | | | | | | | | | | | | | |  
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

RESULT 15

W38851 ID W38851 standard; peptide; 40 AA.  
AC W38851;  
DT 30-MAR-1998 (first entry)  
DE Delivery peptide used in peptide macromolecule complex.  
KW Delivery peptide, peptide-macromolecule complex; macromolecule delivery;  
KW non-exchangeable lipophilic peptide; disease therapy; cell targeting.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc\_difference 39 /note= "any amino acid"  
PN W09725070-A2.  
PD 17-JUL-1997.  
PF 02-JAN-1997; U00454.  
PR 08-JAN-1996; US-584043.  
PA (BAYLOR ) BAYLOR COLLEGE MEDICINE.  
PI Hauer J, Mims MP, Smith LC, Sparrow JT;  
DR WPI; 97-372622/34.  
PT New lipophilic peptide-macromolecule complexes - used for the  
PT delivery of macromolecules to cells, particularly for gene therapy  
PS Claim 6; Page 74; 106pp; English.  
CC This sequence represents a delivery peptide that can be used in the  
CC peptide-macromolecule complex of the invention. The peptide-macromolecule  
CC complex of the invention is for delivering a macromolecule into a cell,  
CC and comprises a non-exchangeable lipophilic peptide (LP) comprising a  
CC delivery peptide associated with a lipid moiety, where the delivery  
CC peptide portion of the LP is complexed to the macromolecule. The  
CC complexes can be used for the delivery of macromolecules such as nucleic  
CC acids, proteins, oligonucleotides, lipids or carbohydrates. They can be  
CC used to treat diseases by enhancing delivery of specific nucleic acid to  
CC the appropriate targeted cells. They can also be used to create  
CC transformed cells as well as transgenic animals for assessing human  
CC disease in an animal model. They can also be used for livestock  
CC agricultural purposes. The complex is capable of transporting the  
CC macromolecule in a stable and condensed state and releasing the molecule  
CC into the cellular interior. The complex can bind with a cell surface  
CC receptor, lyse an endosome and target the nucleus of the cell.  
SQ Sequence 40 AA;

Query Match 100.0%; Score 150; DB 1; Length 40;  
Best Local Similarity 100.0%; Pred. No. 9.6e-10;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30  
| | | | | | | | | | | | | | | | | | | | |  
Db 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30

Search completed: September 7, 1999, 20:37:00  
Job time: 18449 sec





```
; Sequence 11, Application US/07814964
; Patent No. 5359047
; GENERAL INFORMATION:
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pfl, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellett, Patti
; APPLICANT: Essigmann, John M.
; APPLICANT: Lippard, Stephen J.
; TITLE OF INVENTION: DNA Structure Specific Recognition
; TITLE OF INVENTION: Protein and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,964
; FILING DATE: 19911226
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539,906
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4787AAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 11
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; IMMEDIATE SOURCE:
; CLONE: Drosophila SSRP (predicted)
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 458..507
; OTHER INFORMATION: /label= Acidic
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 518..547
; OTHER INFORMATION: /label= Basic I
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 547..620
; OTHER INFORMATION: /label= HMG-box
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label= Basic II
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label= Mixed Charge
; US-07-814-964-11
```

```
Query Match 63.3%; Score 95; DB 1; Length 723;
Best Local Similarity 60.0%; Pred. NO. 0.0023;
Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Oy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 30
Db 519 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 548

RESULT 6
US-08-258-442-11
; Sequence 11, Application US/08258442
; Patent No. 5670621
; GENERAL INFORMATION:
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pfl, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellett, Patti
; APPLICANT: Essigmann, John M.
; APPLICANT: Lippard, Stephen J.
; TITLE OF INVENTION: DNA Structure Specific Recognition
; TITLE OF INVENTION: Protein and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,442
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539,906
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4787AAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 11:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; IMMEDIATE SOURCE:
; CLONE: Drosophila SSRP (predicted)
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 458..507
; OTHER INFORMATION: /label= Acidic
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 518..547
; OTHER INFORMATION: /label= Basic I
; FEATURE:
; NAME/KEY: Domain
```

```
; LOCATION: 547..620
; OTHER INFORMATION: /label= HMG-box
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label= Basic II
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label= Mixed Charge
US-08-258-442-11
```

```
Query Match 63.3%; Score 95; DB 1; Length 723;
Best Local Similarity 60.0%; Pred. No. 0.0023;
Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30
Db 519 KKKKKKKKKKKKKKKKKKKKKKKKK 548
```

```
RESULT 7
US-08-328-809-6
; Sequence 6, Application US/08328809
; Patent No. 5705334
; GENERAL INFORMATION:
; APPLICANT: Lippard, Stephen J.
; APPLICANT: Essigmann, John M.
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pil, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellett, Patti
; TITLE OF INVENTION: Uses For DNA Structure-Specific
; TITLE OF INVENTION: Recognition Proteins
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Administrator, Testa, Hurwitz & Thibault
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
```

```
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,809
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fenton, Gillian M.
REGISTRATION NUMBER: 36,508
REFERENCE/DOCKET NUMBER: MIT-023 (5473/24)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7000
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
IMMEDIATE SOURCE:
CLONE: Drosophila SSRP (predicted)
FEATURE:
; NAME/KEY: Domain
```

```
; LOCATION: 458..507
; OTHER INFORMATION: /label= Acidic
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 518..547
; OTHER INFORMATION: /label= Basic I
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 547..620
; OTHER INFORMATION: /label= HMG
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label= Basic II
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label= Mixed Charge
US-08-328-809-6
```

```
Query Match 63.3%; Score 95; DB 1; Length 723;
Best Local Similarity 60.0%; Pred. No. 0.0023;
Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKK 30
Db 519 KKKKKKKKKKKKKKKKKKKKKKKKK 548
```

```
RESULT 8
PCT-US92-11107-11
; Sequence 11, Application PC/TUS92111107
; GENERAL INFORMATION:
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pil, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellett, Patti
; APPLICANT: Essigmann, John M.
; APPLICANT: Lippard, Stephen J.
; TITLE OF INVENTION: DNA Structure Specific Recognition
; TITLE OF INVENTION: Protein and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/11107
; FILING DATE: 19921218
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539,906
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4787AAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 723 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Drosophila melanogaster  
IMMEDIATE SOURCE:  
CLONE: Drosophila SSRP (predicted)  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 458..507  
OTHER INFORMATION: /label= Acidic  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 518..547  
OTHER INFORMATION: /label= Basic I  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 547..620  
OTHER INFORMATION: /label= HMG-box  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 632..649  
OTHER INFORMATION: /label= Basic II  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 657..723  
OTHER INFORMATION: /label= Mixed Charge  
PCT-US92-11107-11

Query Match 63.3%; Score 95; DB 3; Length 723;  
Best Local Similarity 60.0%; Pred. No. 0.0023;  
Matches 18; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30  
DB 519 KKKKKKKKKKKKKKKKKKKKKKKKK 548

RESULT 9  
US-08-152-488-8  
Sequence 8, Application US/08152488  
Patent No. 5534619  
GENERAL INFORMATION:  
APPLICANT: Wakefield, Thomas W.  
APPLICANT: Andrews, Philip C.  
APPLICANT: Stanley, James C.  
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND  
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN  
TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Benita J. Rohm, Esq.  
STREET: 512 Springfield Avenue  
CITY: Cranford  
STATE: New Jersey  
COUNTRY: United States of America  
ZIP: 07016-1811  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6; ASCII (DOS)Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/152,488  
FILING DATE: 12-NOV-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08069  
FILING DATE: 14-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.

REGISTRATION NUMBER: 28,664  
REFERENCE/DOCKET NUMBER: RM-7WG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-276-3344  
TELEFAX: 908-276-5543  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: N/A  
PUBLICATION INFORMATION:  
AUTHORS: N/A  
TITLE: N/A  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: PCT/US92/08069  
FILING DATE: 14-AUG-1993  
US-08-152-488-8

Query Match 60.0%; Score 90; DB 1; Length 32;  
Best Local Similarity 66.7%; Pred. No. 0.00085;  
Matches 20; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30  
DB 2 KKKSSSKPKKKKKKKKKKKKKKKKK 31

RESULT 10  
US-08-303-025-8  
Sequence 8, Application US/08303025  
Patent No. 5614494  
GENERAL INFORMATION:  
APPLICANT: Wakefield, Thomas W.  
APPLICANT: Andrews, Philip C.  
APPLICANT: Stanley, James C.  
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND  
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN  
TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Benita J. Rohm, Esq.  
STREET: 150 West Jefferson, Suite 2500  
CITY: Detroit  
STATE: Michigan  
COUNTRY: United States of America  
ZIP: 48226-4415  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS v.6.22  
SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,025  
FILING DATE: 08-SEPT-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06829  
FILING DATE: 14-AUG-1992  
APPLICATION NUMBER: US 08/152,488  
FILING DATE: 12-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.  
REFERENCE/DOCKET NUMBER: 7WH-060548-00231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 313-496-7622  
TELEFAX: 313-496-8454  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
;
US-08-303-025-8

Query Match 60.0%; Score 90; DB 1; Length 32;
Best Local Similarity 56.7%; Pred. No. 0.00085;
Matches 20; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||| | |||| | ||||| |||
Db 2 KKKSSSPVKKKKKPKVKSKKKKGKKK 31

RESULT 11
US-08-677-304-8
; Sequence 8, Application US/08677304
; Patent No. 5721212
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,304
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 5721212 Relevant
; TOPOLOGY: No. 5721212 Relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
;

;
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
;
US-08-677-304-8

Query Match 60.0%; Score 90; DB 1; Length 32;
Best Local Similarity 66.7%; Pred. No. 0.00085;
Matches 20; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKK 30
   |||| | |||| | ||||| |||
Db 2 KKKSSSPVKKKKKPKVKSKKKKGKKK 31

RESULT 12
US-08-769-211-1
; Sequence 1, Application US/08769211
; Patent No. 5830852
; GENERAL INFORMATION:
; APPLICANT: Thatcher et al.
; TITLE OF INVENTION: Improved Pharmaceutical Compositions
; TITLE OF INVENTION: Insulin-Receptor Mediated Gene Therapy.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Inc.
; STREET: 75 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1807
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,211
; FILING DATE: 18-Dec-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kathleen M. Williams
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3255/02514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: xaa at position 20 represents
; OTHER INFORMATION: S-acetimidomethyl-Cys or S-pyridyl-Cys.
;
US-08-769-211-1

Query Match 60.0%; Score 90; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKK 18
   ||||| ||||| |||||
Db 2 KKKKKKKKKKKKKKKKKKK 19
```



RESULT 14  
US-08-257-073-5  
; Sequence 5, Application US/08257073  
; Patent NO. 5766597  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: de Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue, 25th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 10036

```

1 GENERAL INFORMATION:
2
3 APPLICANT: Wakefield, Thomas W.
4
5 APPLICANT: Andrews, Philip C.
6
7 APPLICANT: Stanley, James C.
8
9 TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
10
11 TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
12
13 TITLE OF INVENTION: ANTICOAGULATION REVERSAL
14
15 NUMBER OF SEQUENCES: 13
16
17 CORRESPONDENCE ADDRESS:
18
19 ADDRESSEE: Benita J. Rohm, Esq.
20
21 STREET: 512 Springfield Avenue
22
23 CITY: Cranford
24
25 STATE: New Jersey
26
27 COUNTRY: United States of America
28
29 ZIP: 07016-1811
30
31 COMPUTER READABLE FORM:
32
33 MEDIUM TYPE: Floppy disk
34
35 COMPUTER: IBM PC compatible
36
37 OPERATING SYSTEM: MS-DOS
38
39 SOFTWARE: wordperfect 6; ASCII (DOS)text
40
41 CURRENT APPLICATION DATA:
42
43 APPLICATION NUMBER: US/08/152,488
44
45 FILING DATE: 12-NOV-1993
46

```

CLASSIFICATION: 514  
PRIOR APPLICATION DATA: PCT/US92/08069  
APPLICATION NUMBER: 14-AUG-1993  
FILING DATE: 14-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.  
REGISTRATION NUMBER: 28,664  
REFERENCE/DOCKET NUMBER: RM-7WG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-276-3344  
TELEFAX: 908-276-5543  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: N/A  
PUBLICATION INFORMATION:  
AUTHORS: N/A  
TITLE: N/A  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: PCT/US92/08069  
FILING DATE: 14-AUG-1993  
US-08-152-488-7

Query Match 56.0%; Score 84; DB 1; Length 29;  
Best Local Similarity 71.4%; Pred. No. 0.0029;  
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKK 28  
| | | | | | | | | | | | | | | | | |  
Db 2 KKKGGKKKKGGKKKKGGKKKKGGKKKK 29

Search completed: September 7, 1999, 22:38:29  
Job time: 7913 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2170	100.0	434	1	W21591	Antibiotic potenti
2	354	16.3	1178	1	W30763	Mannose-1-phosphat
3	319	14.7	412	1	W03626	Human thyrotropin
4	318	14.7	1588	1	R45605	Human thyrotropin
5	318	14.7	1663	1	R46608	Malarial PfEMP3 ep
6	313	14.4	1132	1	R07866	Plasmodium falcipa
7	291.5	13.4	462	1	R97866	Chicken leucocytoz
8	291	13.4	290	1	R07866	Portion of peptide
9	285	13.1	929	1	W97293	An annexin binding
10	272	12.5	1743	1	W98879	Streptococcus pneu
11	267	12.3	360	1	W36627	H. pylori GHPO 175
12	265.5	12.2	649	1	W90005	Human follicle sti
13	265	12.2	564	1	R62654	Expressed antigen
14	262	12.1	643	1	R84568	C3 binding protein
15	260	12.0	631	1	W46444	Trypanosoma cruzi
16	258.5	11.9	765	1	P92275	Cbpa, a choline bi
17	255.5	11.8	765	1	R75915	Human topoisomer
18	251.5	11.6	1484	1	W89721	Variant human DNA
19	245	11.3	154	1	R06445	Canine ribosome re
20	239	11.0	432	1	W93954	Recombinant copoly
21	238	11.0	564	1	R84565	Human regulatory m
22	237.5	10.9	562	1	R70491	Trypanosoma cruzi
23	233	10.7	452	1	W98334	Leucocytozoan prot
24	231.5	10.7	558	1	R27363	H. pylori GHPO 173
25	230	10.6	1946	1	R47273	Sequence of a poly
26	224	10.3	378	1	R14162	Lactobacillus bulg
27	222	10.2	1411	1	W03258	PRP encoded by clo
28	221	10.2	378	1	R11660	Nucleolar/endosoma
29	220.5	10.2	1023	1	R13319	PRP 378. Plant pro
30	219	10.1	392	1	R96808	Partial Human Natu
31	217.5	10.0	1343	1	W31866	Neisseria meningit
32	215.5	9.9	1382	1	W31867	Mouse metastasis-a
33	214	9.9	530	1	W97775	Human metastasis-a
34	210.5	9.7	493	1	R26944	Human lens epithel
35	209	9.6	43	1	W38654	P.falciparum LSA g
36	207	9.5	44	1	W38898	Delivery peptide u
37	207	9.5	43	1	W38817	Delivery peptide u
38	206.5	9.5	532	1	R27362	Delivery peptide u
39	206	9.5	45	1	W24450	Sequence of a poly
40	204	9.4	42	1	W38853	Nucleic acid (NA)
41	202.5	9.3	517	1	R23904	Delivery peptide u
42	202	9.3	316	1	R26941	1-Caldesmon. 1-Cal
43	202	9.3	43	1	W38897	P.falciparum LSA-R
44	202	9.3	43	1	W38897	Delivery peptide u

421 KKKKKKKKKKKKK 43A









CC disorders e.g. AIDS, Addison's disease, adult respiratory distress  
 CC syndrome, allergies, anaemia, asthma, atherosclerosis, bronchitis,  
 CC cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis,  
 CC dermatomyositis, diabetes mellitus, emphysema, erythema nodosum, atrophic  
 CC gastritis, glomerulonephritis, gout, Graves' disease, hyperosinophilia,  
 CC irritable bowel syndrome, lupus erythematosus, multiple sclerosis,  
 CC myasthenia gravis, myocardial or pericardial inflammation,  
 CC osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid  
 CC arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis,  
 CC complications of cancer, haemodialysis, extracorporeal circulation,  
 CC viral, bacterial, fungal, parasitic, protozoal, and helminthic infections  
 CC and trauma. The products can also be used for detection, diagnosis and  
 CC drug screening.  
 SQ Sequence 290 AA;

Query Match 13.4%; Score 291; DB 1; Length 290;  
 Best Local Similarity 25.0%; Pred. No. 1.7e-19;  
 Matches 72; Conservative 76; Mismatches 130; Indels 10; Gaps 1;

QY 1 KKK 60  
 DB 2 KRNSPQIKKKRRARRAKQSFDDNSLELKDSSKNTAKPKVEMVSGSDDDDFNKL 61  
 QY 61 KKK 120  
 DB 62 KKAQKQKSNKKWDGSEEDNSKIKERSRINSNGESGDESDEFLQSRKQKKNQNK 121  
 QY 121 KKK 180  
 DB 122 PGPNIESGNEDDASFKIKTVQAKAEKKERERKRDEERAKLRKEKELETKGK 181  
 QY 181 KKK 230  
 DB 182 KOKESQRFEEETVSKVTVDTGVPASEEKAETPTAEDDNEGDKKKKKKKKKKEE 241  
 QY 231 KKK 278  
 DB 242 KEKEKKGPSRATVKQAEALAKLEBEROKREESRIKLELESK 289

RESULT 9  
 ID W14593 standard; Protein; 929 AA.  
 AC W14593;  
 DT 27-OCT-1997 (first entry)  
 DE Streptococcus pneumoniae PspC.  
 KW PspA; pneumococcal surface protein; vaccine; otitis media;  
 OS Streptococcus pneumoniae strain EF6796.  
 FH Key Location/Qualifiers  
 FT peptide 1..37  
 FT /label= Sig\_peptide  
 FT protein 38..929  
 FT /label= Mat\_protein  
 FT domain 38..637  
 FT /label= Alpha-helix  
 FT region 69..637  
 FT /label= Coiled-coil  
 FT /note= "breaks in 7-residue periodicity of the  
 FT coiled-coil occur at amino acids 136-141,  
 FT 261-304 and 383-387"  
 FT region 41..242  
 FT /label= Repeat\_1  
 FT /note= "alpha-helical repeat region"  
 FT region 332..492  
 FT /label= Repeat\_2  
 FT /note= "alpha-helical repeat region"  
 FT domain 627..689  
 FT /label= Proline-rich  
 FT domain 913..929  
 FT /label= C-terminal  
 PN W03709994-A1.

PD 20-MAR-1997.  
 PF 16-SEP-1996; U14819.  
 PR 15-SEP-1995; US-529055.  
 PA (UABR-) UAB RES FOUND.  
 PI Briles DE, Brooks-Walter A, Crain MJ, Hollingshead S;  
 PI McDaniel JS, Swiatlo E, Tart R, Yother J;  
 WPI: 97-202002/18.  
 DR N-PSDB; T61728.  
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA -  
 PT used in vaccines for protecting animals against S.pneumoniae  
 PT infection  
 PS Disclosure; Fig 13; 296pp; English.  
 CC This sequence comprises the pneumococcal protein surface C (pspc)  
 CC of Streptococcus pneumoniae strain EF6796. The sequence was  
 CC deduced from the pspC gene (T61728). Like pspA, PspC has an  
 CC alpha-helical coiled-coil region, proline-rich central region,  
 CC repeat regions, with a choline binding motif, and a C-terminal  
 CC 17-amino acid tail. The 2 polypeptides share 3 regions of high  
 CC sequence identity. One is a protection-eliciting region present  
 CC within the alpha-helical domain. The others are the proline-rich  
 CC domain and a repeat domain shared with other choline-binding  
 CC proteins and thought to play a role in cell surface association.  
 CC PspC and PspA polypeptides, and their fragments, can be used in  
 CC vaccines to protect against S. pneumoniae infection and hence for  
 CC the prevention of diseases such as otitis media, meningitis,  
 CC bacteraemia and pneumonia.  
 SQ Sequence 929 AA;

Query Match 13.1%; Score 285; DB 1; Length 929;  
 Best Local Similarity 21.2%; Pred. No. 1.2e-18;  
 Matches 92; Conservative 107; Mismatches 234; Indels 0; Gaps 0;

QY 1 KKK 60  
 DB 72 KDVKKLLKVVQHTQNVGLITIKLSEIKKKYLYDLKVVNLSEALTSKTETKELTATFQ 131  
 QY 61 KKK 120  
 DB 132 FKADLTPEPKVAQAKVVEAKKAEQEKDRRNYPTITYKTLELEIAESDVEVK 191  
 QY 121 KKK 180  
 DB 192 AELELVKRAKESQDEKIKQAEAEVSKQAEATRLKIKTDREAKRADAKLEAVEK 251  
 QY 181 KKK 240  
 DB 252 NVATSQDQPKRRAKRGVSGELATPDKENDAKSSDSSVGEETLPSPSLNMANESOTEHR 311  
 QY 241 KKK 300  
 DB 312 KDVEYIKKMLSEIOLDRKHTQNVNLKLSAIKTYLYELSVLAKENSKKEELTSKTA 371  
 QY 301 KKK 360  
 DB 372 ELTAAFEQFKDTLKPEKVAEAEKVEAEKADQKEDRRNYPTNTYKTLELEIAES 431  
 QY 361 KKK 420  
 DB 432 DVKVKAEELVKEANESRNEEKIKQAEKVESKAEATRLKIKTDREAKRADAKLEAVEK 491  
 QY 421 KKKKKKKKKKKKK 433  
 DB 492 EESEKKAEEAKQK 504

RESULT 10  
 ID W98879 standard; Protein; 1743 AA.  
 AC W98879;  
 DT 31-MAR-1999 (first entry)  
 DE H. pylori GHPO 1735 protein.  
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;







[illegible]

Search completed: September 7, 1999, 20:37:03

QY	361	KKK	363	
		I	I	
Db	601	KNK	603	
RESULT	15			
W46444				
ID	W46444	standard:	Protein:	631 AA.
AC	W46444;			
DT	14-MAY-1998	(first entry)		
DE	CbpA, a choline binding protein from Streptococcus pneumoniae.			
KW	Choline binding protein; CBP: fibronectin; adherence; vaccine.			
KW	antibody; passive immunisation; diagnosis; screening; treatment;			
KW	prevention; bacterial infection; enolase.			
OS	Streptococcus pneumoniae.			
PN	WO9741151-A2.			
PD	06-NOV-1997.			
PF	01-MAY-1997;	U07198.		
PR	01-MAY-1996;	US-642250.		
PR	01-MAY-1996;	US-016632.		
	(UYRQ ) UNIV ROCKEFELLER.			
PI	Masure HR, Rosenow CI, Tuomanen E, Wizeman TM;			
PI	WPI: 97-549682/50.			
DR	N-PSDB: V05102.			
PT	New isolated streptococcal choline binding proteins - used to			
PT	develop products for treating or preventing bacterial infection and			
PT	for detection, diagnosis and screening			
PS	Claim 5; Pages 113-116; 142pp; English.			
CC	The present sequence represents CbpA, a choline binding protein			
CC	(CBP) of Streptococcus pneumoniae. CBP reduces the adherence			
CC	of S. pneumoniae to fibronectin. CBPs, or peptides derived from			
CC	them, can be used to prepare vaccines. Antibodies against them			
CC	can be used for passive immunisation, diagnostics, or screening.			
CC	The CBPs can also be used for treating or preventing bacterial			
CC	infection, as well as screening, detection and diagnosis.			
SQ	Sequence	631 AA;		

[illegible]

Search completed: September 7, 1999, 20:37:03

Job time: 18452 sec

---

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2170	100.0	434	1	US-08-097-830E-3	Sequence 3, Appl1	
2	2170	100.0	434	1	US-08-456-112B-3	Sequence 3, Appl1	
3	455	21.0	1988	1	US-08-056-200-94	Sequence 94, Appl1	
4	318	14.7	1588	3	PCR-US93-07261-11	Sequence 11, Appl1	
5	318	14.7	1663	3	PCR-US93-07261-16	Sequence 16, Appl1	
6	252	12.1	643	2	US-08-216-894-8	Sequence 8, Appl1	
7	255.5	11.8	765	2	US-08-663-113-2	Sequence 2, Appl1	
8	241	11.1	564	2	US-08-216-894-2	Sequence 2, Appl1	
9	231.5	10.7	558	1	US-08-285-440-6	Sequence 6, Appl1	
10	231.5	10.7	558	1	US-08-630-349-6	Sequence 6, Appl1	
11	202.5	9.3	532	1	US-08-285-440-5	Sequence 5, Appl1	
12	202.5	9.3	532	1	US-08-630-349-5	Sequence 5, Appl1	
13	201.5	9.3	591	1	US-08-451-995A-21	Sequence 21, Appl1	
14	201.5	9.3	591	2	US-08-451-747-21	Sequence 21, Appl1	
15	190	8.8	572	1	US-07-609-716-78	Sequence 78, Appl1	
16	190	8.8	740	1	US-07-609-716-79	Sequence 79, Appl1	
17	188	8.7	723	1	US-07-814-964-11	Sequence 11, Appl1	
18	188	8.7	723	1	US-08-258-442-11	Sequence 11, Appl1	
19	188	8.7	723	1	US-08-328-809-6	Sequence 6, Appl1	
20	188	8.7	723	3	PCR-US92-11107-11	Sequence 11, Appl1	
21	171	7.9	472	2	US-08-216-894-10	Sequence 10, Appl1	
22	169.5	7.8	1400	1	US-08-080-255-7	Sequence 7, Appl1	
23	169.5	7.8	1400	3	PCR-US93-05857-7	Sequence 7, Appl1	
24	167	7.7	885	2	US-08-533-306A-4	Sequence 4, Appl1	
25	167	7.7	816	2	US-08-533-306A-6	Sequence 6, Appl1	
26	167	7.7	885	2	US-08-742-923A-4	Sequence 4, Appl1	
27	167	7.7	816	2	US-08-742-923A-6	Sequence 6, Appl1	
28	165	7.6	709	1	US-07-814-964-7	Sequence 7, Appl1	
29	165	7.6	648	1	US-08-072-070-2	Sequence 2, Appl1	
30	165	7.6	695	1	US-08-127-499A-23	Sequence 23, Appl1	
31	165	7.6	695	1	US-08-482-847-23	Sequence 23, Appl1	
32	165	7.6	709	1	US-08-258-442-7	Sequence 7, Appl1	
33	165	7.6	619	1	US-08-465-746-2	Sequence 2, Appl1	
34	165	7.6	709	1	US-08-328-809-2	Sequence 2, Appl1	
35	165	7.6	619	1	US-08-214-164-2	Sequence 2, Appl1	
36	165	7.6	648	2	US-08-469-434-2	Sequence 2, Appl1	
37	165	7.6	648	2	US-08-214-223-2	Sequence 2, Appl1	
38	165	7.6	648	2	US-08-467-852A-2	Sequence 2, Appl1	
39	165	7.6	619	2	US-08-467-852A-3	Sequence 3, Appl1	

[illegible]



















DB 461 KLKLAERK 468

Search completed: September 7, 1999, 22:38:31  
Job time: 7915 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	863	39.8	1390	2	AS1364	sperm tail-specific	
2	559	29.8	678	2	A54514	glutamic acid-rich	
3	555.5	25.6	488	2	I46014	cyliclin II - bovin	
4	523.5	24.1	706	2	A45990	junctional sarcopl	
5	516.5	23.8	606	2	A43427	neurofilament trip	
6	510	23.5	1020	1	QF80H	neurofilament trip	
7	502.5	23.2	729	2	S68191	triadin - human	
8	481.5	22.2	667	2	A40713	cyliclin I - bovine	
9	474.5	21.9	1526	2	A45605	mature-parasite-in	
10	459	21.2	344	2	S34153	mst101-1 protein -	
11	455	21.0	1898	1	A45973	trichohyalin - hum	
12	447	20.6	451	2	G70241	hypothetical prote	
13	442.5	20.4	598	2	B40713	cyliclin I - human	
14	441.5	20.3	865	2	B71609	hypothetical prote	
15	440.5	20.3	1192	2	A71623	probable secreted	
16	440.5	20.3	754	2	JC5314	CDC28/cdc2-like k	
17	438	20.2	1052	1	A44937	kinetoplast-associ	
18	436	20.1	1549	1	A40691	trichohyalin - she	
19	425.5	19.6	771	1	A33430	h-caldesmon - chic	
20	408.5	18.8	2500	2	B71609	hypothetical prote	
21	390	18.0	1407	1	S28589	trichohyalin - rab	
22	388	17.9	1909	2	A45592	liver stage antige	
23	383.5	17.7	793	1	JH0628	caldesmon - human	
24	373	17.2	1038	2	JC5497	clausurin - chicke	
25	369	17.0	980	2	E71606	hypothetical prote	
26	361	16.6	1072	2	A37221	neurofilament trip	
27	361	16.6	332	2	B43427	neurofilament prot	
28	354	16.3	1178	2	S78475	mannosylphosphoryl	
29	353.5	16.3	1087	1	QFWSH	neurofilament trip	
30	349	16.1	854	2	S20003	neurofilament trip	
31	342	15.8	2464	1	QRMSPI	microtubule-associ	
32	340	15.7	2364	2	A56577	microtubule-associ	
33	333.5	15.4	749	2	A45294	Balbani ring 2.1	
34	330	15.2	2441	2	D71623	erythrocyte membra	
35	325	15.0	448	2	I37271	cyliclin II - human	
36	323	14.9	494	2	A48133	pre-mRNA splicing	
37	316	14.6	1002	2	S70292	FUN12 protein - ye	
38	314	14.5	829	2	E73666	DNA topoisomerase	
39	309.5	14.3	508	2	E71620	hypothetical prote	

[illegible]









```

Query Match      21.2%; Score 459; DB 2; Length 344;
Best Local Similarity 37.2%; Pred. No. 1.3e-25;
Matches 111; Conservative 66; Mismatches 121; Indels 0; Gaps 0;

yy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 60
   ||| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
yb 47 KKSAGRGDGLKKKCAEAARKKCAEAARKKCAEAARKKCAEAARKKCAEAARK 106

yy 61 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 120
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
yb 107 KCAEAARKKCAEAARKKCAEAARKKCAEAARKKCAEAARKKCAEAARKKCAEA 166

```

**b**

167 AA EKKCAEAAQKKKCBELAKKSQAEEAEKKKCBAEAAEKKEAERAKKEAAE 226

[illegible]

RESULT 11  
45973  
trichohyalin - human

```

;Species: homo sapiens (man)
;Date: 03-May-1994 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1999
;Accession: A45973

```

1993. *J. Biol. Chem.* 268, 12164-12176, 1993

Reference number: A45973; MUID:93280194  
Accession: A45973  
Molecule type: DNA

;Residues: 1-1898 <LEE>  
;Cross-references: GB:L09190; NID:g292835; PID:g292836  
;Note: authors translated the codon AGG for residue 1714 as Pro

Covalent modifications to this protein include conversion of arginine to citrulline;  
;Genetics:

DB: 136223; OMIM: 190370  
Map position: 1q21-1q21

Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat  
49-81/Domain: calmodulin repeat homology <EF2>

Query Match	21.0%	Score 455;	DB 1;	Length 1898;
Best Local Similarity	4.6%	pred No	5	8e-25;

Matches	20;	Conservative	349;	Mismatches	64;	Indels	0;	Gaps	0
---------	-----	--------------	------	------------	-----	--------	----	------	---

```

dbb 254 RKEEKLQEEPPQRORELQEEELRLKLERQELRRERQEEQQQQLRREQQLRRQEEE 313

```

314 RREQEERREQEERREQQLRREQEERREQQLRREQEERREQQLRREQEERR 373

[illegible]

```

181 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 240
      :::: :::: :::: :::: :::: :::: | :::: :::: :::: :::: :::: :::: + . . . . !

```

434 NEQQLRRQQLRRQEERHEQKHQERREORLKRQEERRDLWKREEETERHEQERRKQ 493



```
Db 112 KEDNNNNNGTKOIEKKNKINKSLLHRQNELNLQSGKNFQDINKNKGKQDISNSNAEN 171
Qy 61 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 120
Db 172 KDVKEGVLEEKKEEKISDDHVEENKSSDDHVKVEENKSSDDHKKIE 231
Qy 121 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 169
Db 232 EVKVEEHEDEEEDKKEKSENKNKNKNDENDEDEDEDEDEDEDEDEDEDEDEDEDE 291
Qy 170 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 229
Db 292 DDDKEDTKHLEEEENIEEFSDKKNKNGKNDTKKESKDTKEKSDIEKEKSKDK 351
Qy 230 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 289
Db 352 EKEKSKDEKKGDKKEKSKDIEKEKDKDIEKEKSDTAKEKDKDIEKEKSKDM 411
Qy 290 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 349
Db 412 EKLKNQONDEKKDDNEKKNDKQDIHDDNDNDNDEEIEBENDEDEDEDEDEDEDE 471
Qy 350 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 409
Db 472 GKNGNENGENSGNENGENENENENENENENENENENENENENENENENENENEN 531
Qy 410 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 432
Db 532 IKEIENVTNANKENYKINKNSE 554

RESULT 15
A71623
probable secreted protein PFB0115w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: A71623
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, E.V.;
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: A71623
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1192 <GAR>
A:Cross-references: GB:AE001373; GB:AE001362; MID:g3845097; PID:g3845099; TIGR:PFB0115w
A:Experimental source: Clone 3D7
C:Genetics:
A:Gene: PFB0115w

Query Match 20.38; Score 440.5; DB 2; Length 1192;
Best Local Similarity 22.6%; Pred. No. 4.3e-24;
Matches 113; Conservative 142; Mismatches 179; Indels 65; Gaps 4;

Qy 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 53
Db 252 KGKTTTWMKNNRDKKSHKRRKTKQNYKYKENQNIENHIPOSKYQRIELDDNG 311
Qy 53 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 98
Db 312 KELASHKNIKEKGIEKTDTTNIADIKKKEERTNDEKKNQIQLVKQVQLIKVGEET 371
Qy 99 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 124
Db 372 KDKEDKEGTDDEEDTDDEEDTDDEEDTDDEEDTDDEEDTDDEEDTDDEEDTDDE 431
Qy 125 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 184
Db 432 EEELEDEEKESEKDESEKDESEKDESEKDESEKDESEKDESEKDESEKDESEKDE 491
Qy 185 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 235
```

```
Db 492 VDEKKEGYEGTDDDEEDKEEEDDEETKVEEKKTEKDEEGTDYEDTDDSDKDETKV 551
Qy 236 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 295
Db 552 EEKTERDEEETEDDEKETEVEKAKTEKDEEGTDYEDTDDSDKDVETEVEETDAEDKEE 611
Qy 296 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 355
Db 612 NEEGTDDDEEKVETDLDDQEEDEEDKEDDKEKDEDDKEDDKEKDEDDKEDKEDDK 671
Qy 356 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 415
Db 672 EDDKEDDKEDKEDNKEKEDNKEKEDDKEKEDDKEKEDDKEKEDNKEKEDNKEKED 731
Qy 416 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 434
Db 732 EDDKEKHDKHVRRIKKKMK 750
```

Search completed: September 7, 1999, 23:06:16  
Job time: 2464 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:49:59 ; Search time 71.87 Seconds  
(without alignments)  
170.703 Million cell updates/sec

Title: US-09-124-280A-3  
Perfect score: 2170  
Sequence: 1 KKKKKKKKKKKKKKKKKKKK.....KKKKKKKKKKKKKKKKKKKK 434

Scoring table: BLOSUM62

Searched: 77977 seqs; 28268293 residues

Database : SwissProt 37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	863	39.8	1391	1	MS12_DROHY	Q08696 drosophila
2	559	25.8	678	1	GARP_PLAFF	P13816 plasmodium
3	555.5	25.6	488	1	CY12_BOVIN	Q28092 bos taurus
4	510	23.5	1020	1	MLH_HUMAN	P12036 homo sapien
5	506	23.3	633	1	FLH_TETH	P40631 tetrahymena
6	481.5	22.2	667	1	CY11_BOVIN	Q35662 bos taurus
7	459	21.2	344	1	MS11_DROHY	Q08695 drosophila
8	455	21.0	1898	1	TRYH_HUMAN	Q07283 homo sapien
9	442.5	20.4	598	1	CY11_HUMAN	Q35663 homo sapien
10	436	20.1	1349	1	TRYH_SHEEP	P22793 ovis aries
11	432	19.9	1220	1	IF2P_HUMAN	O60841 homo sapien
12	425.5	19.6	771	1	CALD_CHICK	P12957 gallus gall
13	390	18.0	1407	1	TRYH_RABIT	Q37709 oryctolagus
14	383.5	17.7	793	1	CALD_HUMAN	Q05682 homo sapien
15	362	16.7	831	1	NFH_RAT	P16884 rattus norv
16	355	16.4	2468	1	MAPB_HUMAN	P48821 homo sapien
17	354	16.3	1178	1	MNN4_YEAST	P36044 saccharomyc
18	353.5	16.3	1087	1	NFH_MOUSE	P19246 mus musculus
19	342	15.8	2464	1	MAPB_MOUSE	P14873 mus musculus
20	325	15.0	348	1	CY12_HUMAN	Q14093 homo sapien
21	323	14.9	494	1	SFR4_HUMAN	Q08170 homo sapien
22	316	14.6	1002	1	IF2P_YEAST	P39730 saccharomyc
23	314	14.5	829	1	TOPI_XENLA	P41512 xenopus lae
24	309	14.2	1641	1	RA54_HUMAN	P45100 homo sapien
25	296	13.6	877	1	INCE_CHICK	P53352 gallus gall
26	287.5	13.2	843	1	CY11_BRUMA	Q27450 buglia mala
27	281	12.9	1395	1	SP41_YEAST	P38904 saccharomyc
28	275.5	12.7	915	1	NFM_HUMAN	P07197 homo sapien
29	267	12.3	1240	1	YNJ1_YEAST	P53935 saccharomyc
30	265.5	12.2	434	1	YK12_YEAST	P36080 saccharomyc
31	263	12.1	260	1	H11_VOLCA	Q08864 volvox cart
32	263	12.1	248	1	H1_PARAN	P02256 parechinus
33	261	12.0	1085	1	YAF4_SCHPO	Q09863 schizosacch
34	260	12.0	1403	1	YDF3_SCHPO	Q10475 schizosacch
35	259	11.9	1079	1	IF2P_SCHPO	Q10251 schizosacch
36	258.5	11.9	767	1	TOPI1_CRIGO	Q07050 cricetus
37	257	11.8	848	1	NFM_MOUSE	P08553 mus musculus
38	256	11.8	845	1	NFM_RAT	P12839 rattus norv
39	255.5	11.8	765	1	TOPI1_HUMAN	P11387 homo sapien
40	255	11.8	805	1	IF2_AQUAE	O67825 aquifex aeo
41	253.5	11.7	644	1	NFM_RABIT	P54938 oryctolagus
42	253	11.7	721	1	ENP1_TORCA	P14400 torpedo cal
43	250.5	11.5	767	1	TOPI1_MOUSE	Q04750 mus musculus

Query Match 39.8%; Score 863; DB 1; Length 1391;  
Best Local Similarity 43.7%; Pred. No. 4.8e-50;  
Matches 208; Conservative 94; Mismatches 132; Indels 4

Best local similarity	43.74	Fixed. no.	4.06	50	
Matches	208	Conservative	94	Mismatches	132
				Indels	42
				Gaps	7

## ALIGNMENTS

```

RESULT 1
ID MST2_DROHY STANDARD; PRT; 1391 AA.
AC Q08696;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE AXONEME-ASSOCIATED PROTEIN MST101(2).
GN MST101(2).
OS DROSOPHILA HYDEI (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEN;
OC DROSOPHILIDAE; DROSOPHILA.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE: 94200512.
RA NERSEN J., BUNEMANN H., HEINLEIN U.A.;
RT "The drosophila hydei gene Dmst101(1) encodes a testis-specific,
RT repetitive, axoneme-associated protein with differential abundance in
RT y chromosomal deletion mutant files.";
RL DEV. BIOL. 162:414-425(1994).
CC -1- FUNCTION: POSSIBLE STRUCTURAL ROLE IN THE SPERM TAIL.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DOMAIN: THE PREDOMINANT STRUCTURE IS ALPHA-HELICAL.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: X73481; G313202; -
DR PIR: S34154; S34154.
DR FLYBASE; FBgn001l816; Dhyd\mst101.
DR HSSP; P01032; 1CSA.
KW SPERM; REPEAT; MULTIGENE FAMILY.
FT FT 331 1283
FT REPEAT 331 378
FT REPEAT 379 426
FT REPEAT 427 474
FT REPEAT 475 522
FT REPEAT 523 570
FT REPEAT 571 618
FT REPEAT 619 666
FT REPEAT 667 714
FT REPEAT 715 762
FT REPEAT 763 808
FT REPEAT 809 856
FT REPEAT 857 904
FT REPEAT 905 952
FT REPEAT 953 995
FT REPEAT 996 1043
FT REPEAT 1044 1091
FT REPEAT 1092 1139
FT REPEAT 1140 1187
FT REPEAT 1188 1235
FT REPEAT 1236 1283
SQ SEQUENCE 1391 AA; 159000 MW; F557E63F CRC32;
20 X 48 AA TANDEM REPEATS OF [ELTQ]-
K(2)-[KQ]C-X(15)-C-X(15)-C-X(11).
1.
2.
3.
4.
5.
6.
7.
8.
9.
10 (INCOMPLETE).
11.
12.
13.
14 (INCOMPLETE).
15.
16.
17.
18.
19.
20.

```

Query Match 39.8%; Score 863; DB 1; Length 1391;  
Best Local Similarity 43.7%; Pred. No. 4.8e-50;  
Matches 208; Conservative 94; Mismatches 132; Indels 4

Best local similarity	43.7%	Fixed. no. 4.0e 20,							
Matches	208;	Conservative	94;	Mismatches	132;	Indels	42;	Gaps	7;













[illegible]

ITS OWN CALCIUM-DEPENDENT POSTSYNTHETIC PROCESSING DURING TERMINAL DIFFERENTIATION.

-1- SUBUNIT: HOMODIMER (PROBABLE).

-1- TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS THE INNER ROOT SHEAT (IRS) OF HAIR FOLLICLES AND MEDULLA, AND IN THE EPITHELIA OF THE TONGUE, HOOF AND RUMEN.

-1- DOMAIN: CONSISTS OF NINE DOMAINS. DOMAIN 1 CONTAINS TWO EF-HAND CALCIUM-BINDING DOMAINS. DOMAINS 2-4, 6, AND 8 ARE ALMOST ENTIRELY ALPHA-HELICAL, CONFIGURED AS A SERIES OF PEPTIDE REPEATS OF VARYING REGULARITY, AND ARE THOUGHT TO FORM A SINGLE-STRANDED ALPHA-HELICAL ROD STABILIZED BY IONIC INTERACTIONS. DOMAIN 6 IS THE MOST REGULAR AND MAY BIND KIF DIRECTLY BY IONIC INTERACTIONS. DOMAINS 5 AND 7 ARE LESS WELL ORGANIZED AND MAY INDUCE FOLDS IN THE MOLECULE. DOMAIN 9 CONTAINS THE C-TERMINUS, CONSERVED AMONG DIFFERENT SPECIES.

-1- PTM: KNOWN SUBSTRATE OF TRANSGLUTAMINASE. SOME 200 ARGININES ARE PROBABLY CONVERTED TO CITRULLINES BY PEPTIDYLARGININE DEIMIDASE.

-1- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS MAY BE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.

-1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND MORE SPECIFICALLY TO S-100/CABP LIKE PROTEINS IN THE N-TERMINUS.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation



```
SQ SEQUENCE 1220 AA: 138754 MW: 48E870FC CRC32:

Query Match 19.9% Score 432: DB 1: Length 1220;
Best Local Similarity 24.1% Pred. No. 2e-22;
Matches 111: Conservative 125; Mismatches 196; Indels 28; Gaps 2;

Qy 1 KKKKKKKKKKK-----KKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 42
Db 3 KKKKKKKKKKK-----KKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 62
Qy 43 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 102
Db 63 EGSLEAQGIKADRETAVVAKPTENNEEF7SKDKKKGGKQKQSPDNDSEELDKSK 122
Qy 103 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 162
Db 123 SKRTAKPKVEMYSDDDDDFENKLPKAKGKAQSNKKWGDSEDEDSNKKIKERSINS 182
Qy 163 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 222
Db 183 SGESDESDEFLQSRGOKNKNKPGNIESGNEDDDASFIKITVAQKKAERERKK 242
Qy 223 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 272
Db 243 RDEEKAKLRKLKEELETGCKDQSKQKESQKQFEETVSKSVTVDTGVIPASEKAETP 302
Qy 273 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 332
Db 303 TAAEDNEGDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 362
Qy 333 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 392
Db 363 EERIKLEELAKKEERLEQEKRRERKKQKQKERRKKEGKLLTKSREARAREAT 422
Qy 393 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 432
Db 423 LKLLQAGVEVPSKDSLPRKRIYEDKKKKKKKKKKKKKKKKKKKKKK 462

RESULT 12
CALD_CHICK STANDARD; PRT; 771 AA.
AC P12957; Q0756; Q0761; Q92018; Q99230; Q03698;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CALDESINON (CDM).
GN CALDI OR CAD.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (GIZZARD H-CAD ISOFORM).
RC TISSUE=GIZZARD, AND OVIDUCT;
RX MEDLINE; 89340480.
RA BRYAN J., INAI M., LEE R., MOORE P., COOK R.G., LIN W.-G.;
RT "Cloning and expression of a smooth muscle caldesmon.";
RJ J. BIOL. CHEM. 264:13873-13879(1989).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (GIZZARD H-CAD ISOFORM).
RC TISSUE=GIZZARD;
RX MEDLINE; 90026426.
RA HAYASHI K., KANDA K., KIMIZUKA F., KATO I., SOBUE K.;
RT "Primary structure and functional expression of h-caldesmon complementary DNA.";
RJ BIOCHEM. BIOPHYS. RES. COMMUN. 164:503-511(1989).
RN [3]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (BRAIN L-CAD ISOFORM).
RC TISSUE=BRAIN;
RX MEDLINE; 91093148.
RA HAYASHI K., FUJIO Y., KATO I., SOBUE K.;
RT "Structural and functional relationships between h- and
```

```
1-caldesmons.";
RJ J. BIOL. CHEM. 266:355-361(1991).
RN [4]
RP SEQUENCE FROM N.A. (GIZZARD H-CAD; BRAIN L-CAD AND GIZZARD L-CAD).
RX MEDLINE; 94071934.
RA HARUNA M., HAYASHI K., YANO H., TAKEUCHI O., SOBUE K.;
RT "Common structural and expressional properties of vertebrate caldesmon genes.";
RJ BIOCHEM. BIOPHYS. RES. COMMUN. 197:145-153(1993).
RN [5]
RP SEQUENCE FROM N.A. (GIZZARD L-CAD ISOFORM).
RC TISSUE=GIZZARD;
RX MEDLINE; 92042686.
RA BRYAN J., LEE R.;
RT "Sequence of an avian non-muscle caldesmon.";
RJ J. MUSCLE RES. CELL. MOTIL. 12:372-375(1991).
RN [6]
RP SEQUENCE OF 1-15 FROM N.A. (BRAIN L-CAD ISOFORM).
RC TISSUE=BRAIN;
RX MEDLINE; 94271210.
RA YANO H., HAYASHI K., HARUNA M., SOBUE K.;
RT "Identification of two distinct promoters in the chicken caldesmon gene.";
RJ BIOCHEM. BIOPHYS. RES. COMMUN. 201:618-626(1994).
RN [7]
RP SEQUENCE OF 451-756 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=GIZZARD;
RX MEDLINE; 89273666.
RA HAYASHI K., YAMADA S., KANDA K., KIMIZUKA F., KATO I., SOBUE K.;
RT "35 kDa fragment of h-caldesmon conserves two consensus sequences of the tropomyosin-binding domain in troponin T.";
RJ BIOCHEM. BIOPHYS. RES. COMMUN. 161:38-45(1989).
RN [8]
RP SEQUENCE OF 498-525.
RX MEDLINE; 88293484.
RA MORNET D., AUDEMAR E., DERANCOURT J.;
RT "Identification of a 15 kilodalton actin binding region on gizzard caldesmon probed by chemical cross-linking.";
RJ BIOCHEM. BIOPHYS. RES. COMMUN. 154:564-571(1988).
RN [9]
RP PHOSPHORYLATION SITES.
RX MEDLINE; 92041815.
RA MAK A.S., CARPENTER M., SMILLIE L.B., WANG J.H.;
RT "Phosphorylation of caldesmon by p34cdc2 kinase. Identification of phosphorylation sites.";
RJ J. BIOL. CHEM. 266:19971-19975(1991).
RN [10]
RP FUNCTION: ACTIN- AND MYOSIN-BINDING PROTEIN IMPLICATED IN THE REGULATION OF ACTOMYOSIN INTERACTIONS IN SMOOTH MUSCLE AND NONMUSCLE CELLS (COULD ACT AS A BRIDGE BETWEEN MYOSIN AND ACTIN FILAMENTS). STIMULATES ACTIN BINDING OF TROPOMYOSIN WHICH INCREASES THE STABILIZATION OF ACTIN FILAMENT STRUCTURE. IN MUSCLE TISSUES, INHIBITS THE ACTOMYOSIN ATPASE BY BINDING TO F-ACTIN. THIS INHIBITION IS ATTENUATED BY CALCIUM-CALMODULIN AND IS POTENTIATED BY TROPOMYOSIN. INTERACTIONS WITH ACTIN, MYOSIN, TWO MOLECULES OF TROPOMYOSIN AND WITH CALMODULIN. ALSO PLAY AN ESSENTIAL ROLE DURING CELLULAR MITOSIS AND RECEPTOR CAPPING.
RN [11]
RP SUBCELLULAR LOCATION: ON THIN FILAMENTS IN SMOOTH MUSCLE AND ON STRESS FIBERS IN FIBROBLASTS (NONMUSCLE) (BY SIMILARITY).
RN [12]
RP TISSUE SPECIFICITY: HIGH-MOLECULAR-WEIGHT CALDESINON (H-CALDESINON) IS PREDOMINANTLY EXPRESSED IN SMOOTH MUSCLES, WHEREAS LOW-MOLECULAR-WEIGHT CALDESINON (L-CALDESINON) IS WIDELY DISTRIBUTED IN NON-MUSCLE TISSUES AND CELLS. NOT EXPRESSED IN SKELETAL MUSCLE OR HEART.
RN [13]
RP DOMAIN: THE N-TERMINAL PART SEEMS TO BE A MYOSIN/CALMODULIN-BINDING DOMAIN, AND THE C-TERMINAL A TROPOMYOSIN/ACTIN/CALMODULIN-BINDING DOMAIN. THESE TWO DOMAINS ARE SEPARATED BY A CENTRAL HELICAL REGION IN THE MUSCLE FORMS.
RN [14]
RP PTM: IN NON-MUSCLE CELLS, PHOSPHORYLATION BY CDC2 DURING MITOSIS CAUSES CALDESINON TO DISSOCIATE FROM MICROFILAMENTS.
RN [15]
RP PHOSPHORYLATION REDUCES CALDESINON BINDING TO ACTIN, MYOSIN, AND CALMODULIN AS WELL AS ITS INHIBITION OF ACTOMYOSIN ATPASE ACTIVITY. PHOSPHORYLATION ALSO OCCURS IN BOTH QUIESCENT AND DIVIDING SMOOTH MUSCLE CELLS WITH SIMILAR EFFECTS ON THE
```







[illegible]





**THIS PAGE BLANK (USPTO)**

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	730.5	33.7	505	5	Q24898	Q24898	ensis minor
2	705	32.5	1701	5	Q61164	Q61164	plasmodium
3	573.5	26.4	1661	5	Q06166	Q06166	plasmodium
4	565	26.0	390	5	Q18401	Q18401	caenorhabdi
5	537	24.7	1359	5	Q02061	Q02061	caenorhabdi
6	523.5	24.1	705	6	Q28820	Q28820	oryctolagus
7	518.5	23.9	6632	5	Q17362	Q17362	caenorhabdi
8	518.5	23.9	6642	5	Q01761	Q01761	caenorhabdi
9	516.5	23.8	606	6	Q28687	Q28687	oryctolagus
10	502.5	23.2	729	4	Q13061	Q13061	homo sapien
11	498	22.9	511	6	Q28688	Q28688	oryctolagus
12	480	22.1	312	5	Q91570	Q91570	caenorhabdi
13	474.5	21.9	1510	5	Q25920	Q25920	plasmodium
14	472	21.8	894	5	Q18462	Q18462	caenorhabdi
15	465	21.4	460	2	Q92102	Q92102	borrelia bu
16	463	21.3	776	5	Q15793	Q15793	p strain dd
17	447	20.6	451	2	Q05070	Q05070	borrelia bu
18	445.5	20.5	1877	5	Q62235	Q62235	caenorhabdi
19	441.5	20.3	3488	5	Q91257	Q91257	caenorhabdi
20	441.5	20.3	565	5	Q96229	Q96229	plasmodium
21	440.5	20.3	754	4	Q13427	Q13427	homo sapien
22	440.5	20.3	1192	4	Q96127	Q96127	plasmodium
23	438	20.2	1052	5	Q26938	Q26938	trypanosoma
24	434	20.0	1220	4	Q95805	Q95805	homo sapien
25	428.5	19.7	335	5	Q07619	Q07619	caenorhabdi
26	423	19.5	699	10	Q92069	Q92069	arabidopsi
27	408.5	18.8	2500	5	Q96223	Q96223	plasmodium
28	407	18.7	752	11	Q55035	Q55035	rattus norv
29	398	17.9	1909	5	Q25893	Q25893	plasmodium

[illegible]



Db	1265	QEIIMNEDVNEKDTANKDKVIEQEKSEKSEVKEKEDTVNKDKVIGQEWIIIEVEVEEVKKR	1324
Qy	388	KKKKKKKKKKK-----KKKKKKKKK-----KKKKKKKKKKKKKKK	423
Db	1325	VKKRNKNKNKDNVIOEIMNEDVNEKDTESDKRMIGKEVIIIEVKEEVKKRVNKEVKKR	1384
Qy	424	KKKKKKKKKKK	434
Db	1385	VNRNKNKNKK	1395

```

RESULT      4
Q18401
ID          PRELIMINARY:      PRT:      390 AA.
AC          Q18401;
DT          01-NOV-1996 (TrEMBLrel. 01, Created)
DT          01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT          01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE          COSMID C33G8.
DE          C33G8.2.
OS          Caenorhabditis elegans.
OC          Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC          Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=BRISTOL N2;
RX          MEDLINE: 94150718.
RA          WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA          BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA          CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA          GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA          JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA          LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA          PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA          SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA          THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSON R.,
RA          WATSON A., WEINSTOCK L., WILKINSON-SPOAT J., WOHLDMAN P.:
RT          "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT          elegans.";
RL          Nature 368:32-38(1994).
RN          [2]
RP          SEQUENCE FROM N.A.
RC          STRAIN=BRISTOL N2;
RA          GATTUNG S., WU X.;
RL          Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN          [3]
RP          SEQUENCE FROM N.A.
RC          STRAIN=BRISTOL N2;
RA          WATERSTON R.;
RL          Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR          EMBL; U53154; AAC25848.1; -.
SO          SEQUENCE 390 AA: 46670 MW: 79RA8247 CRC32:

```

Query Match	26.0%	Score 565;	DB 5;	Length 390;
Best Local Similarity	37.5%	Pred. No. 2.9e-34;		
Matches 128;	Conservative	67;	Mismatches 146;	Indels 0;
				Gaps 0;

[illegible]

Query Match	24.7%	Score 537;	DB 5;	Length 1359;
Best Local Similarity	28.1%	Pred. No. 5.1e-32;		
Matches 107:	Conservative 110;	Mismatches 164;	Indels 0;	Gaps 0;

[illegible]









```
FT NON_TER      511      511
SQ SEQUENCE     511 AA:   55057 MW: F77F2CA5 CRC32:

Query Match                22.9% Score 498 DB 6 Length 511;
Best Local Similarity    27.8%; Pred. No. 1.7e-29;
Matches 120 Conservative 92 Mismatches 219 Indels 0 Gaps 0;

QY 1 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 60
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 81 KSPDKASPAEVKSAPAEAKSPEAKSPAFAEAKSPEAKSPEAKSPVKEAKSPE 140
QY 61 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 120
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 141 KAKSPVEVKSPAFAEAKSPEAKSPVKEEAAPSFAEAKSPEAKSPEAKSPEKA 200
QY 121 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 180
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 201 EAKSPEAKSPVEVKSPAFAEAKSPEAKSPVKEEAAPSFAEAKSPEAKSPEKA 260
QY 181 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 240
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 261 KSPVKEAAKSPVKEAASPVAEAKSPEAKSPEAKSPEAKSPVKEAASPVAEAKSPE 320
QY 241 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 300
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 321 KAKSPEAKSPVKEEAAPSFAEAKSPEAKSPVKEEAAPSFAEAKSPEAKSPEKETPKK 380
QY 301 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 360
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 381 EEKVKEPPKPQAEBETAPPKVERDSRDEAPKPAEKPAYEKPRESFAEAKKDEADKK 440
QY 361 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 420
Dy : | : | : | : | : | : | : | : | : | : | : | : | :
Db 441 TAAPAKVGEAAEPKETEVAKEPEDAKAKEPSKTEKEPEPXPKEETAAPVPKNAK 500
QY 421 KKKKKKKKKKK 431
Dy : | : | : |
Db 501 EAARKPEPK 511


RESULT 12
ID P91570 PRELIMINARY; PRG 312 AA.
AC AC P91570;
DT 01-MAY-1997 (Tremblrel_03, Created)
DT 01-MAY-1997 (Tremblrel_03, Last sequence update)
DT 01-NOV-1998 (Tremblrel_08, Last annotation update)
DE COSMID ZK354.
GN ZK354.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdittia; Rhabdittida;
OC Rhabdittina; Rhabditoidea; Rhabdittidae; Peloderinae; Caenorhabdittis.
RN [1]
SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2.
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2.
```

[illegible]



Db 265 TSLNDLKKQIESAYNFKKYYKSMEXE 291

Search completed: September 7, 1999, 20:34:30  
Job time: 19726 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:03 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-4

Perfect score: 55

Sequence: 1 KDKDKDKDKD 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	100.0	10	1 W21592	Antibiotic potenti
2	43	78.2	35	1 R87195	KEKE-motif of huma
3	43	78.2	660	1 R96970	Bacterial transfer
4	43	78.2	660	1 R77885	Bacterial transfer
5	43	78.2	36	1 R77950	Antigenic Tbp2 pep
6	43	78.2	35	1 R77951	Antigenic Tbp2 pep
7	43	78.2	660	1 R77887	Bacterial transfer
8	43	78.2	644	1 R77883	Bacterial transfer
9	43	78.2	324	1 W11860	Human proteosome s
10	43	78.2	644	1 W08958	Amino acid sequenc
11	43	78.2	660	1 W08960	Amino acid sequenc
12	43	78.2	660	1 W08962	Amino acid sequenc
13	43	78.2	36	1 W46143	Predicted antigeni
14	43	78.2	35	1 W46144	Predicted antigeni
15	43	78.2	523	1 W43011	Truncated transfer
16	43	78.2	529	1 W43012	Truncated transfer
17	43	78.2	547	1 W43013	Truncated transfer
18	43	78.2	647	1 W43014	Truncated transfer
19	43	78.2	660	1 W53045	H. influenzae stra
20	43	78.2	36	1 W54102	Tbp2 antigenic pep
21	43	78.2	35	1 W54103	Tbp2 antigenic pep
22	43	78.2	644	1 W53043	H. influenzae stra
23	43	78.2	660	1 W53047	H. influenzae stra
24	42	76.4	258	1 W32617	Orion yellow dwarf
25	41	74.5	765	1 P922175	Human topoisomeras
26	41	74.5	1189	1 R56496	TATA-binding prote
27	41	74.5	765	1 R75915	Variant human DNA
28	40.5	73.6	1213	1 W06086	Drosophila TATA-bi
29	40.5	73.6	1213	1 W25029	TATA-binding prote
30	40	72.7	11	1 R62185	Ul snRNP 70K prote
31	40	72.7	11	1 R62134	Ul snRNP 70K prote
32	40	72.7	614	1 R82630	70K autoantigen, p
33	40	72.7	286	1 R97417	Papaya ringspot vi
34	40	72.7	436	1 W03662	Human 70K Ul snRNP
35	40	72.7	1311	1 W52197	Precis coenia (but
36	40	72.7	1311	1 W72971	Precis coenia patc
37	39.5	71.8	157	1 R20094	Pfs16 surface prot
38	39	70.9	631	1 R77894	Bacterial transfer
39	39	70.9	631	1 W08968	Amino acid sequenc
40	39	70.9	631	1 W54126	H. influenzae stra
41	39	70.9	521	1 W74802	Human secreted pro
42	39	70.9	813	1 W84307	A human CDC28-#46
43	38	69.1	279	1 R94526	Korean Viper Salmo

ALIGNMENTS

RESULT 1

W21592  
ID W21592 standard; peptide; 10 AA.  
AC W21592:  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #4.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 8; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
| | | | | | | | | |  
Db 1 KDKDKDKDKD 10

RESULT 2

R87195  
ID R87195 standard; peptide; 35 AA.  
AC R87195;  
DT 16-MAY-1996 (first entry)  
DE KEKE-motif of human 26 S proteasome subunit S12.  
KW Proteasome; activation; cell-mediated immunity; immunogen; tolerance;  
KW KEKE motif.  
OS Homo sapiens.  
PN WO9527058-A1.  
PD 12-OCT-1995.  
PF 01-APR-1994; U03591.  
PR 01-APR-1994; WO-U03591.  
PA (UTAH) UNIV UTAH.  
PI Realini CA, Rechsteiner MC;  
DR WPI; 95-358633/46.  
PT DNA encoding human proteasome activator - used to elicit  
PT cell-mediated immunity or tolerance to a selected immunogenic  
PT peptide  
PT Disclosure: Page 38; 71pp; English.  
CC R87194-95, R87262-64 and R87266 are KEKE motif-contg.  
CC peptides which can be used in a method for inducing cell-mediated  
CC immunity against or tolerance to specific epitopes using plasmids  
CC encoding a human proteasome activator (pA) and appropriate epitope-  
CC bearing peptides adjacent to presentation marker peptides (lysine and  
CC glutamine rich peptides termed KEKE motifs peptides, that mark

CC adjacent peptides for presentation). The method can be used to elicit  
 CC cell-mediated immunity or tolerance to pathogen-encoded peptides or  
 CC tumour specific antigens.  
 SQ Sequence 35 AA;

Query Match 78.2%; Score 43; DB 1; Length 35;  
 Best Local Similarity 70.0%; Pred. No. 0.3;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 :||:|||||  
 Db 17 EDKEKDKDK 26

RESULT 3  
 R96970  
 ID R96970 standard; Protein: 660 AA.

AC R96970; 1996 (first entry)  
 DT 25-JUL-1996 (first entry)  
 DE Bacterial transferrin receptor Tbp2 chain(H. influenzae strain Eagan).

KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.

OS Haemophilus influenzae strain Eagan.

FH Key Location/Qualifiers

FT domain 3..19

FT domain /label= transmembrane\_domain\_1

FT domain 55..71

FT domain /label= transmembrane\_domain\_2

FT misc\_difference 59..60

FT domain /note= "Ser-Leu in R77885"

FT domain 126..144

FT domain /label= transmembrane\_domain\_3

FT domain 151..171

FT domain /label= transmembrane\_domain\_4

FT domain 193..206

FT domain /label= transmembrane\_domain\_5

FT misc\_difference 295

FT domain /note= "Glu in R77885"

FT misc\_difference 298

FT domain /note= "Phe in R77885"

FT domain 360..375

FT domain /label= transmembrane\_domain\_6

FT domain 386..405

FT domain /label= transmembrane\_domain\_7

FT misc\_difference 402

FT domain /note= "Ile in R77885"

FT domain 476..491

FT domain /label= transmembrane\_domain\_8

FT domain 508..522

FT domain /label= transmembrane\_domain\_9

FT domain 556..574

FT domain /label= transmembrane\_domain\_10

FT domain 594..611

FT domain /label= transmembrane\_domain\_11

FT domain 623..638

FT domain /label= transmembrane\_domain\_12

FN W09513370-Al.

PD 18-MAY-1995.

PF 07-NOV-1994; CA0616.

PR 08-NOV-1993; US-148968.

PR 29-DEC-1993; US-175116.

PA (CONN-) CONNAUGHT LAB LTD.

PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;

CC (claimed) also stated as being Tbp2 from H. influenzae strain Eagan.  
 CC The Tfr operon consists of two genes (Tbp1 and Tbp2) arranged in tandem  
 CC and are transcribed from a single promoter. The Tbp2 protein tends to  
 CC vary in mol. wt. depending on the species whereas the Tbp1 protein tends  
 CC to have a more consistent mol. wt. with some variability across the  
 CC various bacteria which have Tfr genes. H. influenzae Tfr is iron- and/or  
 CC haemin-regulated and a putative fur-binding site has been identified  
 CC upstream of Tbp2. Antibodies blocking this binding site may prevent  
 CC bacterial growth. Fragments of the Tfr (or its genes) are useful in  
 CC vaccines to provide protection against, e.g. bacterial meningitis. An  
 CC advantage of using the Tfr is that it shares homology with Tfr of other  
 CC H. influenzae strains including non-typable strains.  
 SQ Sequence 660 AA;

Query Match 78.2%; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |||:|:|:|  
 Db 461 KDKDKDKDKD 470

RESULT 4

ID R77885  
 AC R77885 standard; Protein: 660 AA.

DT 24-JUL-1996 (revised)

DT 13-JUN-1996 (first entry)

DE Bacterial transferrin receptor Tbp2 chain(H. influenzae strain Eagan).

KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;

KW non-typable strain; Haemophilus influenzae; meningitis.

OS Haemophilus influenzae strain Eagan.

FN W09513370-Al.

PD 18-MAY-1995.

PF 07-NOV-1994; CA0616.

PR 08-NOV-1993; US-148968.

PR 29-DEC-1993; US-175116.

PA (CONN-) CONNAUGHT LAB LTD.

PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;

PI Murdin A, Schryvers A, Yang Y;

DR N-PSDB; Q94443.

DR N-PSDB; Q94443.

PT Nucleic acids encoding Haemophilus transferrin receptor - used to

PT develop prods for detection and in diagnosis, prevention and

PT treatment of Haemophilus infection.

PS Claim 33; Fig 4A-Q; 231pp; English.

CC The present sequence is that of the transferrin receptor (Tfr) Tbp2

CC chain isolated from Haemophilus influenzae type b, strain Eagan. The Tfr

CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and are

CC transcribed from a single promoter. The Tbp2 protein tends to vary in

CC mol. wt. depending on the species whereas the Tbp1 protein tends to have

CC a more consistent mol. wt. with some variability across the various

CC bacteria which have Tfr genes. H. influenzae Tfr is iron- and/or haemin-

CC regulated and a putative fur-binding site has been identified upstream of

CC Tbp2. Antibodies blocking this binding site may prevent bacterial

CC growth. Fragments of the Tfr (or its genes) are useful in vaccines to

CC provide protection against, e.g. bacterial meningitis. An advantage of

CC using the Tfr is that it shares homology with Tfr of other H. influenzae

CC strains including non-typable strains.  
 SQ Sequence 660 AA;

Query Match 78.2%; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |||:|:|:|  
 Db 461 KDKDKDKDKD 470

PT Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Disclosure; Fig 16B'-B''; 231pp; English.

CC The present sequence is that of the transferrin receptor (Tfr) Tbp2  
 CC chain isolated from Haemophilus influenzae type b, strain Eagan. This  
 CC sequence has a number of amino acid anomalies when compared to R77885



RESULT 5  
 ID R77950 standard; Peptide; 36 AA.  
 AC R77950:  
 DT 09-OCT-1996 (first entry)  
 DE Antigenic Tbp2 peptide TB2-M.  
 KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.  
 OS Synthetic.  
 PN W09513370-A1.  
 PD 18-MAY-1995.  
 PF 07-NOV-1994; CA0616.  
 PR 08-NOV-1993; US-148968.  
 PR 29-DEC-1993; US-175116.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;  
 PI Murdin A, Schryvers A, Yang Y;  
 DR WPI: 95-194089/25.  
 PT Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Example 16; Page 72; 231pp; English.  
 CC R77933-969 are predicted antigenic peptides derived from conserved  
 CC regions of the Tbp2 protein from H. influenzae strains Egan, Minna,  
 CC DL63 and non-typable strain PAK12085. The transferrin receptor (TfR)  
 CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and  
 CC which are transcribed from a single promoter. H. influenzae TfR is  
 CC iron- and/or haemin-regulated and a putative fur-binding site has been  
 CC identified upstream of Tbp2. Antibodies blocking this binding site may  
 CC prevent bacterial growth. Fragments of the TfR (or its genes) are useful  
 CC in vaccines to provide protection against, e.g. bacterial meningitis.  
 CC An advantage of using the TfR is that it shares homology with TfR of  
 CC other H. influenzae strains including non-typable strains. According to  
 CC the specification the present sequence shows residues 360-406 of Tbp2  
 CC from the H. influenzae strain Egan.  
 SQ Sequence 36 AA;

Query Match 78.2%; Score 43; DB 1; Length 36;  
 Best Local Similarity 70.0%; Pred. No. 0.31;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 22 KEKEKDKDKE 31

RESULT 6  
 ID R77951 standard; Peptide; 35 AA.  
 AC R77951:  
 DT 09-OCT-1996 (first entry)  
 DE Antigenic Tbp2 peptide TB2-19.  
 KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.  
 OS Synthetic.  
 PN W09513370-A1.  
 PD 18-MAY-1995.  
 PF 07-NOV-1994; CA0616.  
 PR 08-NOV-1993; US-148968.  
 PR 29-DEC-1993; US-175116.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;  
 PI Murdin A, Schryvers A, Yang Y;  
 DR WPI: 95-194089/25.  
 PT Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Example 16; Page 72; 231pp; English.  
 CC R77933-969 are predicted antigenic peptides derived from conserved  
 CC regions of the Tbp2 protein from H. influenzae strains Egan, Minna,  
 CC DL63 and non-typable strain PAK12085. The transferrin receptor (TfR)  
 CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and

CC which are transcribed from a single promoter. H. influenzae TfR is  
 CC iron- and/or haemin-regulated and a putative fur-binding site has been  
 CC identified upstream of Tbp2. Antibodies blocking this binding site may  
 CC prevent bacterial growth. Fragments of the TfR (or its genes) are useful  
 CC in vaccines to provide protection against, e.g. bacterial meningitis.  
 CC An advantage of using the TfR is that it shares homology with TfR of  
 CC other H. influenzae strains including non-typable strains. According to  
 CC the specification the present sequence shows residues 393-428 of Tbp2  
 CC from the H. influenzae strain Egan.  
 SQ Sequence 35 AA;

Query Match 78.2%; Score 43; DB 1; Length 35;  
 Best Local Similarity 70.0%; Pred. No. 0.3;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 1 KEKDKDKDKE 10

RESULT 7  
 ID R77887 standard; Protein; 660 AA.  
 AC R77887:  
 DT 09-OCT-1996 (first entry)  
 DE Bacterial transferrin receptor Tbp2 (H. influenzae strain Minna).  
 KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.  
 KW Haemophilus influenzae strain Minna.  
 PN W09513370-A1.  
 PD 18-MAY-1995.  
 PF 07-NOV-1994; CA0616.  
 PR 08-NOV-1993; US-148968.  
 PR 29-DEC-1993; US-175116.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;  
 PI Murdin A, Schryvers A, Yang Y;  
 DR WPI: 95-194089/25.  
 DR N-PSDB; Q94444.  
 PT Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Claim 33; Fig 5A-Q; 231pp; English.  
 CC The present sequence is that of the transferrin receptor (TfR) Tbp2  
 CC chain isolated from Haemophilus influenzae type b, strain Minna. The TfR  
 CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and are  
 CC transcribed from a single promoter. The Tbp2 protein tends to vary in  
 CC mol. wt. depending on the species whereas the Tbp1 protein tends to have  
 CC a more consistent mol. wt. with some variability across the various  
 CC bacteria which have TfR genes. H. influenzae TfR is iron- and/or haemin-  
 CC regulated and a putative fur-binding site has been identified upstream of  
 CC Tbp2. Antibodies blocking this binding site may prevent bacterial  
 CC growth. Fragments of the TfR (or its genes) are useful in vaccines to  
 CC provide protection against, e.g. bacterial meningitis. An advantage of  
 CC using the TfR is that it shares homology with TfR of other H. influenzae  
 CC strains including non-typable strains.  
 SQ Sequence 660 AA;

Query Match 78.2%; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 461 KDKDKDKDKD 470

RESULT 8  
 ID R77883 standard; Protein; 644 AA.  
 AC R77883:  
 DT 09-OCT-1996 (first entry)  
 DE Bacterial transferrin receptor Tbp2 (H. influenzae strain Minna).  
 KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis.  
 KW Haemophilus influenzae strain Minna.  
 PN W09513370-A1.  
 PD 18-MAY-1995.  
 PF 07-NOV-1994; CA0616.  
 PR 08-NOV-1993; US-148968.  
 PR 29-DEC-1993; US-175116.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;  
 PI Murdin A, Schryvers A, Yang Y;  
 DR WPI: 95-194089/25.  
 DR N-PSDB; Q94444.  
 PT Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Claim 33; Fig 5A-Q; 231pp; English.  
 CC The present sequence is that of the transferrin receptor (TfR) Tbp2  
 CC chain isolated from Haemophilus influenzae type b, strain Minna. The TfR  
 CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and are  
 CC transcribed from a single promoter. The Tbp2 protein tends to vary in  
 CC mol. wt. depending on the species whereas the Tbp1 protein tends to have  
 CC a more consistent mol. wt. with some variability across the various  
 CC bacteria which have TfR genes. H. influenzae TfR is iron- and/or haemin-  
 CC regulated and a putative fur-binding site has been identified upstream of  
 CC Tbp2. Antibodies blocking this binding site may prevent bacterial  
 CC growth. Fragments of the TfR (or its genes) are useful in vaccines to  
 CC provide protection against, e.g. bacterial meningitis. An advantage of  
 CC using the TfR is that it shares homology with TfR of other H. influenzae  
 CC strains including non-typable strains.  
 SQ Sequence 644 AA;

Wed Sep 8 09:44:54 1999

09-OCT-1996 (first entry)  
 DE Bacterial transferrin receptor Tbp2 (H. influenzae strain DL63).  
 KW Tbp1; Tbp2; transferrin receptor operon; vaccine; antigen;  
 KW non-typable strain; Haemophilus influenzae; meningitis; ss.  
 OS Haemophilus influenzae strain DL63  
 FH Key Location/Qualifiers  
 FT peptide 1..25  
 FT /label= signal\_peptide  
 FT protein 26..644  
 FT /label= mature\_protein  
 FT /note= "Tbp2 chain of Tfr"  
 FT W09513370-A1.  
 PN 18-MAY-1995.  
 PD 07-NOV-1994; CA0616.  
 PF 08-NOV-1993; US-148968.  
 PR 29-DEC-1993; US-175116.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness R, Klein M, Loosmore S;  
 PI Murdin A, Schryvers A, Yang Y;  
 DR N-PSDB; Q94442.  
 DR Nucleic acids encoding Haemophilus transferrin receptor - used to  
 PT develop prods for detection and in diagnosis, prevention and  
 PT treatment of Haemophilus infection.  
 PS Claim 3; Fig 3A-Q; 231pp; English.  
 CC The present sequence is that of the transferrin receptor (Tfr) Tbp2  
 CC chain isolated from Haemophilus influenzae type b, strain DL63. The Tfr  
 CC operon consists of two genes (Tbp1 and Tbp2) arranged in tandem and are  
 CC transcribed from a single promoter. The Tbp2 protein tends to vary in  
 CC mol. wt. depending on the species whereas the Tbp1 protein tends to have  
 CC a more consistent mol. wt. with some variability across the various  
 CC bacteria which have Tfr genes. H. influenzae Tfr is iron- and/or haemin-  
 CC regulated and a putative fur-binding site has been identified upstream of  
 CC Tbp2. Antibodies blocking this binding site may prevent bacterial  
 CC growth. Fragments of the Tfr (or its genes) are useful in vaccines to  
 CC provide protection against, e.g. bacterial meningitis. An advantage of  
 CC using the Tfr is that it shares homology with Tfr of other H. influenzae  
 CC strains including non-typable strains.  
 SQ Sequence 644 AA;

Query Match 78.2%; Score 43; DB 1; Length 644;  
 Best Local Similarity 70.0%; Pred. No. 5.6;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 454 KEKDKDKDK 463

RESULT 9  
 W11860  
 ID W11860 standard; Protein; 324 AA.  
 AC W11860;  
 DT 16-APR-1997 (first entry)  
 DE Human proteosome subunit P40.  
 KW Human proteosome P40 subunit; protease; diagnosis; treatment;  
 KW immune disease; heart tissue.  
 OS Homo sapiens.  
 PN J08308567-A.  
 PD 26-NOV-1996.  
 PF 19-MAY-1995; 121484.  
 PR 19-MAY-1995; JP-121484.  
 PA (SUME) SUMITOMO ELECTRIC IND CO.  
 DR WPI; 97-059690/06.  
 DR N-PSDB; T61366.  
 PT Human proteosome sub-unit P40 and corresponding gene - used in the  
 PT treatment of various immune diseases  
 PS Claim 1; Page 11-12; 15pp; Japanese.  
 CC This sequence represents human proteosome P40 subunit. The proteosome  
 CC subunit has protease activity and may be used in a kit to diagnose  
 CC and treat various immune diseases. Expression of DNA encoding the  
 CC P40 subunit was found to be high in heart tissue and low in brain,

CC lung, liver, pancreas and placenta.  
 SQ Sequence 324 AA;

Query Match 78.2%; Score 43; DB 1; Length 324;  
 Best Local Similarity 70.0%; Pred. No. 2.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 302 EDKDKDKDK 311

RESULT 10  
 W08958  
 ID W08958 standard; Protein; 644 AA.  
 AC W08958;  
 DT 05-MAY-1998 (first entry)  
 DE Amino acid sequence of the transferrin chain Tbp2.  
 KW Transferrin receptor; Haemophilus influenzae type b strain DL63;  
 KW iron; human transferrin; iron source; antibody; bacterial growth;  
 KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
 OS Haemophilus influenzae.  
 FH Key Location/Qualifiers  
 FT Peptide 1..17  
 FT /note= "signal peptide"  
 FT W09640929-A2.  
 PN 19-DEC-1996; CA0399.  
 PD 07-JUN-1996; US-649518.  
 PR 17-MAY-1996; US-483577.  
 PR 07-JUN-1995; US-483577.  
 PA (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-owen S, Harkness RE, Klein MH, Loosmore SM,  
 PI Murdin AD, Schryvers AB, Yang Y;  
 DR WPI; 97-052329/05.  
 DR N-PSDB; T49500.  
 DR Haemophilus truncated transferrin receptor protein analogue, Tbp2  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus tfr  
 PT antibodies  
 PS Claim 5; Fig 3A-Q; 228pp; English.  
 CC The present sequence represents the transferrin receptor chain Tbp2 of  
 CC Haemophilus influenzae type b, strain DL63. The bacterial transferrin  
 CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a  
 CC non-encapsulated or non-typable bacterium responsible for a wide range  
 CC of human diseases. Iron is an essential nutrient for the growth of these  
 CC bacteria, and they can utilise human transferrin as a source of iron.  
 CC Antibodies which block the access of the transferrin receptor to  
 CC its iron source prevent bacterial growth. The transferrin receptor, or  
 CC fragments, therefore, are good vaccine candidates. The full length Tbp2  
 CC protein is produced in low amounts in Escherichia coli. However, the  
 CC yield can be enhanced by truncation of the 3' end of the gene. An  
 CC immunogenic composition comprising (or encoding) the immunogenic  
 CC truncated analogue can be used to induce protection against a disease  
 CC caused by a bacterial pathogen that produces the transferrin receptor.  
 CC The immunogenic truncated analogue is also useful as an antigen in  
 CC immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor  
 CC genes.  
 SQ Sequence 644 AA;

Query Match 78.2%; Score 43; DB 1; Length 644;  
 Best Local Similarity 70.0%; Pred. No. 5.6;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |:|:|:|:|:  
 Db 454 KEKDKDKDK 463

RESULT 11

W08960  
 ID W08960 standard; Protein; 660 AA.  
 AC W08960;  
 DT 05-MAY-1998 (first entry)  
 DE Amino acid sequence of the transferrin chain Tbp2.  
 KW Transferrin receptor; Haemophilus influenzae type b;  
 KW iron; human transferrin; iron source; antibody; bacterial growth;  
 KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
 OS Haemophilus influenzae.  
 PN W09640929-A2.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; CA0399.  
 PR 17-MAY-1996; US-649518.  
 PR 07-JUN-1995; US-483577.  
 PR (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,  
 PI Murdin AD, Schryvers AB, Yang Y;  
 DR WPI: 97-052329/05.  
 N-PSDB: 749501.  
 DT Haemophilus truncated transferrin receptor protein analogue, Tbp2 -  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus Tfr  
 PT antibodies  
 PS Claim 5; Fig 4A-Q: 228pp; English.  
 CC The present sequence represents the transferrin receptor chain Tbp2 of  
 CC Haemophilus influenzae type b, strain Minna. The bacterial transferrin  
 CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a  
 CC non-encapsulated or non-typable bacterium responsible for a wide range  
 CC of human diseases. Iron is an essential nutrient for the growth of these  
 CC bacteria, and they can utilise human transferrin as a source of iron.  
 CC Antibodies which block the access of the transferrin receptor to  
 CC its iron source prevent bacterial growth. The transferrin receptor, or  
 CC fragments, therefore, are good vaccine candidates. The full length Tbp2  
 CC protein is produced in low amounts in Escherichia coli. However, the  
 CC yield can be enhanced by truncation of the 3' end of the gene. An  
 CC immunogenic composition comprising (or encoding) the immunogenic  
 CC truncated analogue can be used to induce protection against a disease  
 CC caused by a bacterial pathogen that produces the transferrin receptor.  
 CC The immunogenic truncated analogue is also useful as an antigen in  
 CC immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor  
 CC genes.  
 SQ Sequence 660 AA;  
 Query Match 78.28; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KDKDKDKDKD 10  
 Db 461 KDKDKDKDKD 470  
 |||:|:|:|  
 W08962  
 ID W08962 standard; Protein; 660 AA.  
 AC W08962; (first entry)  
 DE Amino acid sequence of the transferrin chain Tbp2.  
 KW Transferrin receptor; Haemophilus influenzae type b;  
 KW iron; human transferrin; iron source; antibody; bacterial growth;  
 KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
 OS Haemophilus influenzae.  
 PN W09640929-A2.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; CA0399.  
 PR 17-MAY-1996; US-649518.  
 PR 07-JUN-1995; US-483577.  
 PR (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,  
 PI Murdin AD, Schryvers AB, Yang Y;  
 DR WPI: 97-052329/05.  
 N-PSDB: 749501.  
 DT Haemophilus truncated transferrin receptor protein analogue, Tbp2 -  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus Tfr  
 PT antibodies  
 PS Claim 5; Fig 4A-Q: 228pp; English.  
 CC The present sequence represents the transferrin receptor chain Tbp2 of  
 CC Haemophilus influenzae type b, strain Egan. The bacterial transferrin  
 CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a  
 CC non-encapsulated or non-typable bacterium responsible for a wide range  
 CC of human diseases. Iron is an essential nutrient for the growth of these  
 CC bacteria, and they can utilise human transferrin as a source of iron.  
 CC Antibodies which block the access of the transferrin receptor to  
 CC its iron source prevent bacterial growth. The transferrin receptor, or  
 CC fragments, therefore, are good vaccine candidates. The full length Tbp2  
 CC protein is produced in low amounts in Escherichia coli. However, the  
 CC yield can be enhanced by truncation of the 3' end of the gene. An  
 CC immunogenic composition comprising (or encoding) the immunogenic  
 CC truncated analogue can be used to induce protection against a disease  
 CC caused by a bacterial pathogen that produces the transferrin receptor.  
 CC The immunogenic truncated analogue is also useful as an antigen in  
 CC immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor  
 CC genes.  
 SQ Sequence 660 AA;  
 Query Match 78.28; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KDKDKDKDKD 10  
 Db 461 KDKDKDKDKD 470  
 |||:|:|:|  
 W08962  
 ID W08962 standard; Protein; 660 AA.  
 AC W08962; (first entry)  
 DE Amino acid sequence of the transferrin chain Tbp2.  
 KW Transferrin receptor; Haemophilus influenzae type b;  
 KW iron; human transferrin; iron source; antibody; bacterial growth;  
 KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
 OS Haemophilus influenzae.  
 PN W09640929-A2.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; CA0399.  
 PR 17-MAY-1996; US-649518.  
 PR 07-JUN-1995; US-483577.  
 PR (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,  
 PI Murdin AD, Schryvers AB, Yang Y;  
 DR WPI: 97-052329/05.  
 N-PSDB: 749501.  
 DT Haemophilus truncated transferrin receptor protein analogue, Tbp2 -  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus Tfr  
 PT antibodies  
 PS Claim 5; Fig 4A-Q: 228pp; English.  
 CC The present sequence represents the transferrin receptor chain Tbp2 of  
 CC Haemophilus influenzae type b, strain Egan. The bacterial transferrin  
 CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a  
 CC non-encapsulated or non-typable bacterium responsible for a wide range  
 CC of human diseases. Iron is an essential nutrient for the growth of these  
 CC bacteria, and they can utilise human transferrin as a source of iron.  
 CC Antibodies which block the access of the transferrin receptor to  
 CC its iron source prevent bacterial growth. The transferrin receptor, or  
 CC fragments, therefore, are good vaccine candidates. The full length Tbp2  
 CC protein is produced in low amounts in Escherichia coli. However, the  
 CC yield can be enhanced by truncation of the 3' end of the gene. An  
 CC immunogenic composition comprising (or encoding) the immunogenic  
 CC truncated analogue can be used to induce protection against a disease  
 CC caused by a bacterial pathogen that produces the transferrin receptor.  
 CC The immunogenic truncated analogue is also useful as an antigen in  
 CC immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor  
 CC genes.  
 SQ Sequence 660 AA;  
 Query Match 78.28; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KDKDKDKDKD 10  
 Db 461 KDKDKDKDKD 470  
 |||:|:|:|

DR WPI: 97-052329/05.  
 PT Haemophilus truncated transferrin receptor protein analogue, Tbp2 -  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus Tfr  
 PT antibodies  
 PS Claim 5; Fig 5A-Q: 228pp; English.  
 CC The present sequence represents the transferrin receptor chain Tbp2 of  
 CC Haemophilus influenzae type b, strain Minna. The bacterial transferrin  
 CC receptor is composed of 2 chains, Tbp1 and Tbp2. H. influenzae is a  
 CC non-encapsulated or non-typable bacterium responsible for a wide range  
 CC of human diseases. Iron is an essential nutrient for the growth of these  
 CC bacteria, and they can utilise human transferrin as a source of iron.  
 CC Antibodies which block the access of the transferrin receptor to  
 CC its iron source prevent bacterial growth. The transferrin receptor, or  
 CC fragments, therefore, are good vaccine candidates. The full length Tbp2  
 CC protein is produced in low amounts in Escherichia coli. However, the  
 CC yield can be enhanced by truncation of the 3' end of the gene. An  
 CC immunogenic composition comprising (or encoding) the immunogenic  
 CC truncated analogue can be used to induce protection against a disease  
 CC caused by a bacterial pathogen that produces the transferrin receptor.  
 CC The immunogenic truncated analogue is also useful as an antigen in  
 CC immunoassays for the detection of Haemophilus transferrin receptor  
 CC antibodies, while the nucleic acid molecule can be used as a  
 CC hybridisation probe for the detection of other transferrin receptor  
 CC genes.  
 SQ Sequence 660 AA;  
 Query Match 78.28; Score 43; DB 1; Length 660;  
 Best Local Similarity 70.0%; Pred. No. 5.8;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KDKDKDKDKD 10  
 Db 461 KDKDKDKDKD 470  
 |||:|:|:|  
 W46143  
 ID W46143 standard; Protein; 36 AA.  
 AC W46143; (first entry)  
 DT 05-MAY-1998 (first entry)  
 DE Predicted antigenic Tbp1 peptide TBP2-M.  
 KW Transferrin receptor; Haemophilus influenzae type b;  
 KW iron; human transferrin; iron source; antibody; bacterial growth;  
 KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
 OS Synthetic.  
 OS Haemophilus influenzae.  
 PN W09640929-A2.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; CA0399.  
 PR 17-MAY-1996; US-649518.  
 PR 07-JUN-1995; US-483577.  
 PR (CONN-) CONNAUGHT LAB LTD.  
 PI Chong P, Gray-Owen S, Harkness RE, Klein MH, Loosmore SM,  
 PI Murdin AD, Schryvers AB, Yang Y;  
 DR WPI: 97-052329/05.  
 DT Haemophilus truncated transferrin receptor protein analogue, Tbp2 -  
 PT used to induce protection against disease caused by transferrin  
 PT producing pathogens, or as antigen to detect Haemophilus Tfr  
 PT antibodies  
 PS Example 16; Page 70; 228pp; English.  
 CC Protein of Haemophilus influenzae type b. Tbp2 is part of the  
 CC transferrin receptor, of which Tbp1 is also a subunit. The deduced amino  
 CC acid sequences of Tbp1 and Tbp2 were compared, and regions of  
 CC conservation identified. The above peptides are derived from these  
 CC regions, the present peptide being derived from residues 360-406.  
 CC Iron is an essential nutrient for the growth of these bacteria, and they  
 CC can utilise human transferrin as a source of iron. Antibodies which  
 CC block the access of the transferrin receptor to its iron source prevent  
 CC bacterial growth. The transferrin receptor, or fragments, therefore,  
 CC are good vaccine candidates. An immunogenic composition comprising

CC (or encoding) the immunogenic truncated analogue can be used  
CC to induce protection against a disease caused by a bacterial pathogen  
CC that produces the transferrin receptor. The immunogenic truncated  
CC analogue is also useful as an antigen in immunoassays for the detection  
CC of Haemophilus transferrin receptor antibodies, while the nucleic acid  
CC molecule can be used as a hybridisation probe for the detection of other  
CC transferrin receptor genes.  
SQ Sequence 36 AA;

Query Match 78.2%; Score 43; DB 1; Length 36;  
Best Local Similarity 70.0%; Pred. No. 0.3;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
I:|:|:|:|:|:  
Db 22 KEKDKDKDK 31

RESULT 14

W46144 ID W46144 standard; Protein: 35 AA.  
AC W46144;  
DT 05-MAY-1998 (first entry)  
DE Predicted antigenic Tbp1 peptide TBP2-19  
KW Transferrin receptor; Haemophilus influenzae type b;  
KW iron; human transferrin; iron source; antibody; Bacterial growth;  
KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
OS Synthetic.  
OS Haemophilus influenzae.  
PN W09640929-A2.  
PD 19-DEC-1996.  
PF 07-JUN-1996; CA0399.  
PR 17-MAY-1996; US-649518.  
PR 07-JUN-1995; US-483577.  
PA (CONN-) CONNAUGHT LAB LTD.  
PI Chong P. Gray-Owen S. Harkness RE, Klein MH, Loosmore SM,  
PI Murdin AD, Schryvers AB, Yang Y;  
DR WPI: 97-052329/05.  
PT Haemophilus truncated transferrin receptor protein analogue, Tbp2  
PT used to induce protection against disease caused by transferrin  
PT producing pathogens, or as antigen to detect Haemophilus Tfr  
PT antibodies.  
PS Example 16: Page 70: 228pp: English.  
CC W46126-62 are predicted antigenic peptides derived from the Tbp2  
CC protein of Haemophilus influenzae type b. Tbp2 is part of the  
CC transferrin receptor, of which Tbp1 is also a subunit. The deduced amino  
CC acid sequences of Tbp1 and Tbp2 were compared, and regions of  
CC conservation identified. The above peptides are derived from these  
CC regions, the present peptide being derived from residues 393-428.  
CC Iron is an essential nutrient for the growth of these bacteria, and they  
CC can utilise human transferrin as a source of iron. Antibodies which  
CC block the access of the transferrin receptor to its iron source prevent  
CC bacterial growth. The transferrin receptor, or fragments, therefore,  
CC are good vaccine candidates. An immunogenic composition comprising  
CC (or encoding) the immunogenic truncated analogue can be used  
CC to induce protection against a disease caused by a bacterial pathogen  
CC that produces the transferrin receptor. The immunogenic truncated  
CC analogue is also useful as an antigen in immunoassays for the detection  
CC of Haemophilus transferrin receptor antibodies, while the nucleic acid  
CC molecule can be used as a hybridisation probe for the detection of other  
CC transferrin receptor genes.  
SQ Sequence 35 AA;

Query Match 78.2%; Score 43; DB 1; Length 35;  
Best Local Similarity 70.0%; Pred. No. 0.3;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
I:|:|:|:|:|:  
Db 1 KEKDKDKDK 10

RESULT 15

W43011 ID W43011 standard; Protein: 523 AA.  
AC W43011;  
DT 05-MAY-1998 (first entry)  
DE Truncated transferrin receptor chain Tbp2 protein DS-1461-1-1.  
KW Transferrin receptor; Haemophilus influenzae type b strain Egan;  
KW iron; human transferrin; iron source; antibody; Bacterial growth;  
KW vaccine; immunogenic truncated analogue; antigen; Tbp1; Tbp2.  
OS Synthetic.  
OS Haemophilus influenzae.  
PN W09640929-A2.  
PD 19-DEC-1996.  
PF 07-JUN-1996; CA0399.  
PR 17-MAY-1996; US-649518.  
PR 07-JUN-1995; US-483577.  
PA (CONN-) CONNAUGHT LAB LTD.  
PI Chong P. Gray-Owen S. Harkness RE, Klein MH, Loosmore SM,  
PI Murdin AD, Schryvers AB, Yang Y;  
DR WPI: 97-052329/05.  
PT Haemophilus truncated transferrin receptor protein analogue, Tbp2  
PT used to induce protection against disease caused by transferrin  
PT producing pathogens, or as antigen to detect Haemophilus Tfr  
PT antibodies.  
PS Claim 4; Fig 31A-B: 228pp: English.  
CC W43001-14 represent truncated analogues of the Tbp2 chain of the  
CC transferrin receptor of Haemophilus influenzae type b, strain Egan.  
CC The transferrin receptor is composed of 2 chains, Tbp1 and Tbp2.  
CC H. influenzae is a non-encapsulated or non-typable bacterium responsible  
CC for a wide range of human diseases. Iron is an essential nutrient for  
CC the growth of these bacteria, and they can utilise human transferrin as  
CC a source of iron. Antibodies which block the access of the transferrin  
CC receptor to its iron source prevent bacterial growth. The transferrin  
CC receptor, or fragments, therefore, are good vaccine candidates. The full  
CC length Tbp2 protein is produced in low amounts in Escherichia coli.  
CC However, the yield can be enhanced by truncation of the 3' end of the  
CC gene. An immunogenic composition comprising (or encoding) the  
CC immunogenic truncated analogue can be used to induce protection against  
CC a disease caused by a bacterial pathogen that produces the transferrin  
CC receptor. The immunogenic truncated analogue is also useful as an antigen  
CC in immunoassays for the detection of Haemophilus transferrin receptor  
CC antibodies, while the nucleic acid molecule can be used as a  
CC hybridisation probe for the detection of other transferrin receptor  
CC genes.  
SQ Sequence 523 AA;

Query Match 78.2%; Score 43; DB 1; Length 523;  
Best Local Similarity 70.0%; Pred. No. 4.6;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
I:|:|:|:|:|:  
Db 461 KDKDKDKDKD 470

Search completed: September 7, 1999, 20:37:05  
Job time: 18454 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:31 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-4  
Perfect score: 55  
Sequence: 1 KDKDKDKDKD 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	2	US-08-456-112B-4
2	43	78.2	644	1	US-08-487-890A-6
3	43	78.2	660	1	US-08-487-890A-8
4	43	78.2	660	1	US-08-487-890A-10
5	43	78.2	36	1	US-08-487-890A-66
6	43	78.2	35	1	US-08-487-890A-67
7	41	74.5	765	2	US-08-663-112-2
8	40.5	73.6	1213	1	US-08-188-582-20
9	40.5	73.6	1213	1	US-08-646-715-20
10	40	72.7	11	1	US-08-704-170-34
11	40	72.7	11	1	US-08-704-170-93
12	40	72.7	1311	2	US-08-540-406-4
13	40	72.7	11	3	PCT-US94-02631-34
14	40	72.7	11	3	PCT-US94-02631-93
15	40	72.7	614	3	PCT-US95-03236-21
16	40	72.7	220	3	PCT-US95-03236-36
17	40	72.7	1311	3	PCT-US95-13233-4
18	39.5	71.8	157	2	US-08-450-065-2
19	39.5	71.8	157	2	US-08-450-595-2
20	39	70.9	631	1	US-08-487-890A-111
21	37	67.3	866	1	US-08-100-692-1
22	37	67.3	479	1	US-08-583-318-5
23	37	67.3	633	1	US-08-458-477A-5
24	36	65.5	162	1	US-08-244-557-2
25	36	65.5	232	2	US-08-516-801-4
26	36	65.5	183	3	PCT-US95-03236-37
27	36	65.5	232	3	PCT-US95-06683-4
28	35	63.6	887	1	US-07-867-106-3
29	35	63.6	798	1	US-08-190-802A-64
30	35	63.6	798	1	US-08-190-802A-68
31	35	63.6	10	1	US-08-097-830E-4
32	35	63.6	740	2	US-08-257-073-5
33	35	63.6	594	2	US-08-785-310A-6
34	35	63.6	798	2	US-08-308-818-2
35	35	63.6	325	2	US-08-828-242-4
36	34	61.8	331	2	US-08-828-242-3
37	33	60.0	620	1	US-08-325-071-65
38	33	60.0	776	1	US-08-021-601-2
39	33	60.0	456	1	US-08-021-601-6

Sequence 8, Appli  
Sequence 10, Appli  
Sequence 2, Appli  
Sequence 6, Appli  
Sequence 8, Appli  
Sequence 10, Appli

ALIGNMENTS

RESULT 1  
US-08-456-112B-4  
; Sequence 4, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: POTO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8988  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-4  
Query Match 100.0%; Score 55; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
Db 1 KDKDKDKDKD 10

RESULT 2  
US-08-487-890A-6  
; Sequence 6, Application US/08487890A  
; Patent No. 5708149  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-ping

APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487.890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 644 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-6

Query Match 78.2%; Score 43; DB 1; Length 644;  
Best Local Similarity 70.0%; Pred. No. 2.8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
|:|:|:|:|:|:  
Db 454 KEREKDKKE 463

RESULT 3  
US-08-487-890A-8  
Sequence 8, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487.890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-8

Query Match 78.2%; Score 43; DB 1; Length 660;  
Best Local Similarity 70.0%; Pred. No. 2.9;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
|:|:|:|:|:|:  
Db 461 KDKDKDKDKD 470

RESULT 4  
US-08-487-890A-10  
Sequence 10, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487.890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116

; FILING DATE: 29-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,968  
; FILING DATE: 08-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 660 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-487-890A-10

Query Match 78.2%; Score 43; DB 1; Length 660;  
Best Local Similarity 70.0%; Pred. No. 2.9;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKKDKDKD 10  
Db 461 KDKKDKDKD 470

RESULT 5  
US-08-487-890A-66  
; Sequence 66, Application US/08487890A  
; Patent No. 5708149  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-ping  
; APPLICANT: Murdin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 Unviersity Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,890A  
; FILING DATE: 07-JUN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,116  
; FILING DATE: 29-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,968  
; FILING DATE: 08-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-487-890A-66

Query Match 78.2%; Score 43; DB 1; Length 36;  
Best Local Similarity 70.0%; Pred. No. 0.17;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
Db 22 KEKEKDKKE 31

RESULT 6  
US-08-487-890A-67  
; Sequence 67, Application US/08487890A  
; Patent No. 5708149  
; GENERAL INFORMATION:  
; APPLICANT: Loosmore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-ping  
; APPLICANT: Murdin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sim & McBurney  
; STREET: 6th Floor, 330 Unviersity Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,890A  
; FILING DATE: 07-JUN-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,116  
; FILING DATE: 29-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/148,968  
; FILING DATE: 08-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-487-890A-67

Query Match 78.2%; Score 43; DB 1; Length 35;  
Best Local Similarity 70.0%; Pred. No. 0.17;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKD 10  
|:|||||:|:  
Db 1 KEKDKDKE 10

## RESULT 7

US-08-663-112-2  
; Sequence 2, Application US/08663112  
; Patent No. 5849503  
; GENERAL INFORMATION:  
; APPLICANT: WAGATSUMA, Masako  
; APPLICANT: KURITA, No. 5849503iko  
; TITLE OF INVENTION: MUTANT PROTEINS OF HUMAN DNA  
; TITLE OF INVENTION: TOPOISOMERASE I  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/663,112  
; FILING DATE: 26-NOV-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Einaudi, Carolyn P.  
; REGISTRATION NUMBER: 32,220  
; REFERENCE/DOCKET NUMBER: 06609.1488-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 765 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-663-112-2

Query Match 74.5%; Score 41; DB 2; Length 765;  
Best Local Similarity 70.0%; Pred. No. 6.8;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKD 10  
||| |||:  
Db 78 KDKDKDKE 87

## RESULT 8

US-08-188-582-20  
; Sequence 20, Application US/08188582  
; Patent No. 5534410  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.  
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,582  
; FILING DATE: 28-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1213 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-188-582-20

Query Match 73.6%; Score 40.5; DB 1; Length 1213;  
Best Local Similarity 81.8%; Pred. No. 13;  
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 KDKDKD-KDKD 10  
|||||:||||  
Db 1170 KDKDKDKD 1180

## RESULT 9

US-08-646-715-20  
; Sequence 20, Application US/08646715  
; Patent No. 5637886  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.  
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/08/646,715  
FILING DATE: 09-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/188,582  
FILING DATE: 28-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1213 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-715-20

Query Match 73.68; Score 40.5; DB 1; Length 1213;  
Best Local Similarity 81.8%; Pred. No. 13;  
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 1 KDKDKD-KDKD 10  
Db 1170 KDKDKERKDKD 1180

RESULT 10  
US-08-704-170-34  
Sequence 34 Application US/08704170  
Patent No. 5707626  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 No. 5707626th Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704,170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-34

Query Match 72.7%; Score 40; DB 1; Length 11;  
Best Local Similarity 50.0%; Pred. No. 0.16;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
Db 1 RDRDRDRDRD 10

RESULT 11  
US-08-704-170-93  
Sequence 93 Application US/08704170  
Patent No. 5707626  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 No. 5707626th Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704,170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-93

Query Match 72.7%; Score 40; DB 1; Length 11;  
Best Local Similarity 50.0%; Pred. No. 0.16;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
Db 1 RDRDRDRDRD 10

RESULT 12  
US-08-540-406-4  
Sequence 4, Application US/08540406  
Patent No. 5837538

GENERAL INFORMATION:  
APPLICANT: SCOTT, MATHEW P  
APPLICANT: GOODRICH, LISA V  
APPLICANT: JOHNSON, RONALD L  
TITLE OF INVENTION: Patched Genes and their Use  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/540,406  
FILING DATE: 06-OCT-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: a60190-1  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1311 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-540-406-4

Query Match 72.7%; Score 40; DB 2; Length 1311;  
Best Local Similarity 50.0%; Pred. No. 16;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
Db 1260 RDRDRDRDRD 1269

RESULT 13  
PCT-US94-02631-34  
Sequence 34, Application PC/TUS9402631  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 North Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02631  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-34  
Query Match 72.7%; Score 40; DB 3; Length 11;  
Best Local Similarity 50.0%; Pred. No. 0.16;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KDKDKDKDKD 10  
Db 1 RDRDRDRDRD 10  
RESULT 14  
PCT-US94-02631-93  
Sequence 93, Application PC/TUS9402631  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 North Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02631  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-93

Query Match 72.7%; Score 40; DB 3; Length 11;  
 Best Local Similarity 50.0%; Pred. No. 0.16;  
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
 :|:|:|:|:  
 Db 1 RDRDRDRD 10

RESULT 15  
 PCT-US95-03236-21  
 ; Sequence 21, Application PC/TUS9503236  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Southern California  
 ; TITLE OF INVENTION: Methods to Diagnose and Treat HIV-1  
 ; TITLE OF INVENTION: Infection  
 ; NUMBER OF SEQUENCES: 66  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell and Flores  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/03236  
 ; FILING DATE: 13-MAR-1995  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Imbra, Richard J.  
 ; REGISTRATION NUMBER: 37,643  
 ; REFERENCE/DOCKET NUMBER: FP-SI 1394  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 21:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 614 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 PCT-US95-03236-21

Query Match 72.7%; Score 40; DB 3; Length 614;  
 Best Local Similarity 50.0%; Pred. No. 7.9;  
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
 :|:|:|:|:  
 Db 543 RDRDRDRD 552

Search completed: September 7, 1999, 22:38:32  
 Job time: 7916 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:16 : Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-4

Perfect score: 55

Sequence: 1 KDKDKDKDKD 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	92.7	471	2 S02016	U1 snRNP 70K prote
2	51	92.7	228	2 S55888	CCHH finger protei
3	51	92.7	1997	2 F71607	DNA helicase II BR
4	50	90.9	160	2 A45607	16K surface antige
5	46	83.6	375	2 A40112	MHC-region RD-repe
6	45	81.8	737	2 T02178	hypothetical prote
7	45	81.8	2206	2 G71611	hypothetical prote
8	45	81.8	326	2 I51694	xbmi-1 - African c
9	44	80.0	822	1 TVHUFF	protein-tyrosine k
10	44	80.0	820	1 TVCTFF	protein-tyrosine k
11	44	80.0	820	2 J48347	protein-tyrosine k
12	43	78.2	324	2 JC4154	26S proteasome reg
13	43	78.2	321	2 S65491	26S proteasome reg
14	43	78.2	660	2 S70904	transferrin-bindin
15	43	78.2	577	2 S30237	transcription init
16	42	76.4	767	2 JU0144	DNA topoisomerase
17	42	76.4	1027	2 A56533	iswi protein - fru
18	41	74.5	765	1 I5HUT1	DNA topoisomerase
19	41	74.5	767	2 A49546	DNA topoisomerase
20	41	74.5	2364	2 A44159	spectrin beta-G ch
21	41	74.5	707	2 S29854	spectrin beta chai
22	41	74.5	421	2 S36799	calreticulin precu
23	41	74.5	400	2 S43376	calreticulin, brai
24	41	74.5	1163	2 D64315	type I restriction
25	41	74.5	665	2 B71609	hypothetical prote
26	41	74.5	1192	2 A71623	probable secreted
27	41	74.5	1041	2 H71617	SERA antigen/papai
28	41	74.5	708	2 A38436	mitosis initiation
29	40.5	73.6	1213	2 A54063	TATA-binding prote
30	40	72.7	325	2 JH0189	arginine/aspartate
31	40	72.7	614	2 A25707	U1 snRNP 70K prote
32	40	72.7	382	2 A33640	class III histocom
33	40	72.7	380	2 S36789	gene RD protein -
34	40	72.7	156	2 S03047	U1 snRNP 70K prote
35	40	72.7	541	2 T01394	hypothetical prote
36	40	72.7	1516	2 E71619	RAD2 endonuclease
37	40	72.7	484	2 A40988	54k arginine-rich
38	39.5	71.8	157	2 S30129	macrogamete/sporo
39	39	70.9	813	2 JC5785	ATP-dependent RNA

methy1 coenzyme M  
transferrin-bindin  
cell division cont  
Drosophila translo  
neurotactin - frui  
neurotactin - frui

ALIGNMENTS

RESULT 1

S02016

U1 snRNP 70K protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 10-Sep-1997

C:Accession: S02016

R:Ezerodt, M.; Vignali, R.; Ciliberto, G.; Scherly, D.; Mattaj, I.W.; Philipson, L.

EMBO J. 7, 4311-4321, 1988

A:Title: Structure and expression of a Xenopus gene encoding an snRNP protein (U1 70K

A:Reference number: S02016; MUID:89210819

A:Accession: S02016

A:Molecule type: DNA

A:Residues: 1-471 <EPZ>

A:Cross-references: EMBL:X12430; NID:g65178; PID:g65179

C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleopro  
F:104-174/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 92.7%; Score 51; DB 2; Length 471;

Best Local Similarity 90.0%; Pred. No. 0.45;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10

|||||:||||

Db 275 KDKDKDKDKD 284

RESULT 2

S55888

CCHH finger protein 8 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 25-Apr-1997

C:Accession: S55888

R:Tague, B.W.; Goodman, H.M.

Plant Mol. Biol. 28, 267-279, 1995

A:Title: Characterization of a family of Arabidopsis zinc finger protein cDNAs.

A:Reference number: S55881; MUID:95322589

A:Accession: S55888

A:Molecule type: mRNA

A:Residues: 1-228 <TAG>

A:Cross-references: GB:L39651

C:Keywords: DNA binding; zinc finger

F:84-104/Region: zinc finger CCHH motif

Query Match 92.7%; Score 51; DB 2; Length 228;

Best Local Similarity 90.0%; Pred. No. 0.22;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10

|||||:||||

Db 62 KDKDKDKDKD 71

RESULT 3

F71607

DNA helicase II BRAHMA homolog PF0730w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 21-Nov-1998

C:Accession: F71607

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H

Science 282, 1126-1132, 1998  
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
 A:Reference number: A71600  
 A:Accession: F71607  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1997 <GAR>  
 A:Cross-references: GB:AE001414; GB:AE001362; NID:g3845260; PID:g3845262; TIGR:PF0730W  
 A:Experimental source: clone 3D7  
 C:Genetics:  
 A:Gene: PFB0730W

Query Match 92.7%; Score 51; DB 2; Length 1997;  
 Best Local Similarity 90.0%; Pred. No. 1.9;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |||||

DB 1618 KDKDKDKDKDK 1627

## RESULT 4

A45607  
 16K surface antigen Pvs16 - Plasmodium reichenowi  
 C:Species: Plasmodium reichenowi  
 C:Date: 22-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
 C:Accession: A45607  
 R:Moelans, I.I.; Lal, A.A.; Konings, R.N.; Schoenmakers, J.G.  
 Mol. Biochem. Parasitol. 50, 349-350, 1992  
 A:Title: Sequence of a 16-kilodalton sexual stage and sporozoite surface antigen of Plasmodium falciparum  
 A:Reference number: A45607; MUID:92158015  
 A:Accession: A45607  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-160 <MOE>  
 A:Note: sequence inconsistent with the nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:83676, NCBIIP:83681)  
 C:Keywords: surface antigen

Query Match 90.9%; Score 50; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKDKDKDKD 10  
 |||||

DB 141 DKDKDKDKD 149

## RESULT 5

A40112  
 MHC-region RD-repeat protein - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 10-Sep-1997  
 C:Accession: A40112  
 R:Levi-Strauss, M.; Carroli, M.C.; Steinmetz, M.; Meo, T.  
 Science 240, 201-204, 1988  
 A:Title: A previously undetected MHC gene with an unusual periodic structure.  
 A:Reference number: A40112; MUID:88178091  
 A:Accession: A40112  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-375 <LEV>  
 A:Cross-references: GB:M21332; NID:gl199607; PID:gl199608  
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein  
 F:267-326/Domain: ribonucleoprotein repeat homology <RRM3>

Query Match 83.6%; Score 46; DB 2; Length 375;  
 Best Local Similarity 70.0%; Pred. No. 2;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 :||:|||||

DB 212 RDRDKDKDKD 221

## RESULT 6

T02178  
 hypothetical protein F14M4.8 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 05-Mar-1999  
 C:Accession: T02178  
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;  
 submitted to the EMBL Data Library, September 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.  
 A:Reference number: Z14609

A:Accession: T02178  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-737 <ROU>  
 A:Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522937  
 C:Genetics:  
 A:Map position: 2  
 A:Introns: 56/3; 159/3; 200/1; 237/3; 267/3  
 A:Note: F14M4.8

Query Match 81.8%; Score 45; DB 2; Length 737;  
 Best Local Similarity 80.0%; Pred. No. 5.4;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 :|||||

DB 634 KSQDKDKDKD 643

## RESULT 7

G71611  
 hypothetical protein PFB0560W - malaria parasite (Plasmodium falciparum)  
 C:Species: Plasmodium falciparum  
 C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
 C:Accession: G71611  
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;  
 ; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
 Science 282, 1126-1132, 1998  
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
 A:Reference number: A71600

A:Accession: G71611  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-2206 <GAR>  
 A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PID:g3845217; TIGR:PF056  
 C:Experimental source: clone 3D7  
 C:Genetics:  
 A:Gene: PFB0560W

Query Match 81.8%; Score 45; DB 2; Length 2206;  
 Best Local Similarity 80.0%; Pred. No. 16;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 |||||

DB 666 KDKDKDKDKN 675

## RESULT 8

I51694  
 xhmi-1 - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 10-Jul-1998  
 C:Accession: I51694  
 R:Otte, A.P.; Reijnen, M.J.; Hamer, K.M.; Den Blaauwen, J.L.; Lambrechts, C.; Schonev  
 Mech. Dev. 53, 35-46, 1995

```

A:Title: Polycomb and bmi-1 homologs are expressed in overlapping patterns in Xenopus en
A:Reference number: 151693
A:Accession: F51694
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <OPT>
A:Cross-references: EMBL:U39959; NID:g1086576; PID:g1086577
C:Genetics:
A:Gene: xbm1-1
C:Superfamily: RING finger homology
C:Keywords: zinc
F:14-62/Domain: RING finger homology <RNG>

Query Match      81.8%; Score 45; DB 2; Length 326;
Best Local Similarity 80.0%; Pred. No. 2.4;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 KDKDKDKDKD 10
      |||||:| |
Db      149 KDKDKERSKD 158

RESULT      9
TVHUFF
protein-tyrosine kinase (EC 2.7.1.112) fes/fps - human
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 05-Feb-1999
C:Accession: A24673; A60188
R:Roebroek, A.J.M.; Schalken, J.A.; Verbeek, J.S.; Van den Ouweland, A.M.W.; Onnekink, C.
EMBO J. 4, 2897-2903, 1985
A:Title: The structure of the human c-fes/fps proto-oncogene.
A:Reference number: A24673; MUID:86055727
A:Accession: A24673
A:Molecule type: DNA
A:Residues: 1-822 <ROE>
A:Cross-references: GB:X06292; GB:M14209; GB:M14589; NID:g31348; PID:g31349
R:Alcalay, M.; Antolini, F.; Van de Ven, W.J.; Lanfrancone, L.; Grignani, F.; Pelicci, F.
Oncogene 5, 267-275, 1990
A:Title: Characterization of human and mouse c-fes cDNA clones and identification of the
A:Reference number: A60188; MUID:90191711
A:Accession: A60188
A:Molecule type: mRNA
A:Residues: 1-718; 'S', 720-822 <ALC>
A:Cross-references: GB:X52192; NID:g29890; PID:g29891
A:Note: the authors translated the codon TCC for residue 719 as Leu
C:Genetics:
A:Gene: GDB:FES
A:Cross-references: GDB:119906; OMIM:190030
A:Map position: 15q26.1-15q26.1
A:Introns: 71/3; 129/3; 162/1; 223/2; 269/2; 309/2; 412/3; 440/3; 510/3; 51/3; 56
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho
F:2-822/Product: protein-tyrosine kinase fes/fps #status predicted <MAT>
F:460-545/Domain: SH2 homology <SH2>
F:559-821/Domain: protein kinase homology <KIN>
F:567-575/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:590/Active site: Lys #status predicted

Query Match      80.0%; Score 44; DB 1; Length 822;
Best Local Similarity 80.0%; Pred. No. 8.5;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 KDKDKDKDKD 10
      |||||:| |
Db      161 KDKDKAKD 170

RESULT      10
TVHUFF
protein-tyrosine kinase (EC 2.7.1.112) fes/fps - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 30-Sep-1989 #sequence_revision 30-Jun-1992 #text_change 23-Feb-1997
C:Accession: A27824
R:Roebroek, A.J.M.; Schalken, J.A.; Onnekink, C.; Bloemers, H.P.J.; Van de Ven, W.J.M
J. Virol. 61, 2009-2016, 1987
A:Title: Structure of the feline c-fes/fps proto-oncogene: genesis of a retroviral on
A:Reference number: A27824; MUID:87198954
A:Accession: A27824
A:Molecule type: DNA
A:Residues: 1-820 <ROE>
C:Genetics:
A:Gene: fes/fps
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; proto-oncog
F:458-543/Domain: SH2 homology <SH2>
F:557-819/Domain: protein kinase homology <KIN>
F:565-573/Region: protein kinase ATP-binding motif
F:588/Active site: Lys #status predicted
F:711/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
```

```

TVCTFF
protein-tyrosine kinase (EC 2.7.1.112) fes/fps - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 30-Sep-1989 #sequence_revision 30-Jun-1992 #text_change 23-Feb-1997
C:Accession: A27824
R:Roebroek, A.J.M.; Schalken, J.A.; Onnekink, C.; Bloemers, H.P.J.; Van de Ven, W.J.M
J. Virol. 61, 2009-2016, 1987
A:Title: Structure of the feline c-fes/fps proto-oncogene: genesis of a retroviral on
A:Reference number: A27824; MUID:87198954
A:Accession: A27824
A:Molecule type: DNA
A:Residues: 1-820 <ROE>
C:Genetics:
A:Gene: fes/fps
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; proto-oncog
F:458-543/Domain: SH2 homology <SH2>
F:557-819/Domain: protein kinase homology <KIN>
F:565-573/Region: protein kinase ATP-binding motif
F:588/Active site: Lys #status predicted
F:711/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match      80.0%; Score 44; DB 1; Length 820;
Best Local Similarity 80.0%; Pred. No. 8.5;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 KDKDKDKDKD 10
      |||||:| |
Db      159 KDKDKAKD 168

RESULT      11
I48347
protein-tyrosine kinase-related protein - mouse
N:Alternate names: c-fes protein
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 12-Feb-1999
C:Accession: I48347; JH0112
R:Wilks, A.F.; Kurban, R.R.
Oncogene 3, 289-294, 1988
A:Title: Isolation and structural analysis of murine c-fes cDNA clones.
A:Reference number: I48347; MUID:89083198
A:Accession: I48347
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-820 <RES>
A:Cross-references: EMBL:X12616; NID:g50955; PID:g50956
R:Wilks, A.F.; Kurban, R.R.; Hovens, C.M.; Ralph, S.J.
Gene 85, 67-74, 1989
A:Title: The application of the polymerase chain reaction to cloning members of the p
A:Reference number: JH0112; MUID:90152381
A:Accession: JH0112
A:Molecule type: mRNA
A:Residues: 678-745 <WIL>
A:Experimental source: haemopoietic cell, clone FD15
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:458-543/Domain: SH2 homology <SH2>
F:557-819/Domain: protein kinase homology <KIN>
F:565-573/Region: protein kinase ATP-binding motif

Query Match      80.0%; Score 44; DB 2; Length 820;
Best Local Similarity 80.0%; Pred. No. 8.5;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 KDKDKDKDKD 10
      |||||:| |
Db      161 KDKDKAKD 170

RESULT      12
```

## JC41154

26S proteasome regulatory chain, p40 - human  
N:Alternate names: mov-34 homolog  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1995 #sequence\_revision 10-Nov-1995 #text\_change 08-Sep-1997  
C:Accession: JC41154; PC4037  
R:Tsurumi, C.; DeMartino, G.N.; Slaughter, C.A.; Shimbara, N.; Tanaka, K.  
Biochem. Biophys. Res. Commun. 210, 600-608, 1995  
A:Title: cDNA cloning of p40, a regulatory subunit of the human 26S proteasome, and a homologous protein from Drosophila melanogaster  
A:Reference number: JC41154; MUID:95275316  
A:Accession: JC41154  
A:Molecule type: mRNA  
A:Residues: 1-324 <TSU>  
A:Cross-references: DDBJ:D50063; NID:g971269; PID:d1009401; PID:g971270  
A:Accession: PC4037  
A:Molecule type: protein  
A:Residues: 35-45;94-100;130-143;154-173;181-199;205-214 <TS2>  
C:Comment: This protein is a novel essential regulatory subunit of the human 26S proteasome.  
C:Genetics:  
A:Gene: GDB:PSMD7; MOV34; P40; S12  
A:Cross-references: GDB:127564; OMIM:157970  
A:Map position: 16q22.2-16q22.2  
C:Superfamily: mov-34 protein  
F:286-324/Region: KEKE motif

Query Match 78.2%; Score 43; DB 2; Length 324;  
Best Local Similarity 70.0%; Pred. No. 4.8;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|||:|||||  
DB 302 EDKEKDKDK 311

## RESULT 13

S65491  
26S proteasome regulatory chain 12 - human  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 17-Oct-1997  
R:Dubiel, W.; Ferrell, K.; Dumdey, R.; Standera, S.; Prehn, S.; Rechsteiner, M.  
FEBS Lett. 363, 97-100, 1995  
A:Title: Molecular cloning and expression of subunit 12: a non-MCP and non-ATPase subunit of the 26S proteasome  
A:Reference number: S65491; MUID:95246894  
A:Accession: S65491  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-321 <DUB>  
A:Accession: S65492  
A:Molecule type: protein  
A:Residues: 2-18, 'X', 20-29 <DU2>  
C:Superfamily: mov-34 protein

Query Match 78.2%; Score 43; DB 2; Length 321;  
Best Local Similarity 70.0%; Pred. No. 4.7;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|||:|||||  
DB 302 EDKEKDKDK 311

## RESULT 14

S70904  
transferrin-binding protein 2 precursor - Haemophilus influenzae (strains Egan and Minn)  
C:Species: Haemophilus influenzae  
A:Variety: strains Egan and strain MinnA  
C:Date: 15-Feb-1997 #sequence\_revision 24-Oct-1997 #text\_change 21-Nov-1997  
C:Accession: S70904; S70902  
R:Loosmore, S.M.; Yang, Y.; Coleman, D.C.; Shortreed, J.M.; England, D.M.; Harkness, R.E.  
Mol. Microbiol. 19, 575-586, 1996

A:Title: Cloning and expression of the Haemophilus influenzae transferrin receptor gene  
A:Reference number: S70901  
A:Accession: S70904  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-660 <LOO>  
A:Cross-references: EMBL:U15051; NID:g1223937; PID:g1223938  
A:Experimental source: strain Egan, clone S-4368-3-3, JB-901-5-3  
A:Accession: S70902  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-660 <LO2>  
A:Cross-references: EMBL:U15052; NID:g1223940; PID:g1223941  
A:Experimental source: strain MinnA, clone DS-712-1-3  
C:Genetics:  
A:Gene: tbp2  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-660/Product: transferrin-binding protein 2 #status predicted <MAT>

Query Match 78.2%; Score 43; DB 2; Length 660;  
Best Local Similarity 70.0%; Pred. No. 9.6;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|||:|||||  
DB 461 KDKDKDKDK 470

## RESULT 15

S30237  
transcription initiation factor TFIIF-alpha - fruit fly (Drosophila melanogaster)  
N:Alternate names: transcription factor 5 large chain  
C:Species: Drosophila melanogaster  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Mar-1999  
R:Kephart, D.D.; Price, M.P.; Burton, Z.F.; Finkelstein, A.; Greenblatt, J.; Price, D.  
Nucleic Acids Res. 21, 1319, 1993  
A:Title: Cloning of a Drosophila cDNA with sequence similarity to human transcription factor TFIIF  
A:Reference number: S30237; MUID:93219101  
A:Accession: S30237  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-577 <KEP>  
A:Cross-references: EMBL:X66982; NID:g288190; PID:g288191  
R:Gong, D.W.; Horikoshi, M.; Nakatani, Y.  
Nucleic Acids Res. 21, 1492, 1993  
A:Title: Analysis of cDNA encoding Drosophila transcription initiation factor TFIIF-a  
A:Reference number: S35621; MUID:93219133  
A:Accession: S35621  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-577 <GON>  
A:Cross-references: EMBL:L10331  
R:Gong, D.W.; Horikoshi, M.; Nakatani, Y.  
submitted to the EMBL Data Library, April 1993  
A:Reference number: S65538  
A:Accession: S65538  
A:Molecule type: DNA  
A:Residues: 1-19, 'R', 21-75, 'G', 77-527, 'A', 529-577 <GOW>  
C:Genetics:  
A:Gene: FlyBase:TFIIFaagr  
A:Cross-references: FlyBase:FBgn0010282  
C:Keywords: DNA binding; transcription regulation

Query Match 78.2%; Score 43; DB 2; Length 577;  
Best Local Similarity 70.0%; Pred. No. 8.4;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|||:|||||



Db 425 KDKDKEKE 434

Search completed: September 7, 1999, 23:06:17  
Job time: 2465 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:02 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-4

Perfect score: 55

Sequence: 1 KDKDKDKD 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	92.7	471	1	RUI7_XENLA
2	46	83.6	375	1	RDP_MOUSE
3	44	80.0	820	1	FES_FELCA
4	44	80.0	822	1	FES_HUMAN
5	44	80.0	820	1	FES_MOUSE
6	43	78.2	324	1	PRSC_HUMAN
7	43	78.2	577	1	T2FA_DROME
8	42	76.4	508	1	N060_DROME
9	42	76.4	767	1	TOP1_MOUSE
10	41	74.5	400	1	CRT1_BOVIN
11	41	74.5	421	1	CRT2_BOVIN
12	41	74.5	696	1	FSYA_DROME
13	41	74.5	2364	1	SPCO_HUMAN
14	41	74.5	767	1	TOPI_CRIGR
15	41	74.5	765	1	TOPI_HUMAN
16	41	74.5	1075	1	Y124_METJA
17	40	73.6	1213	1	T2D2_DROME
18	40	72.7	1851	1	CCAL_DROME
19	40	72.7	380	1	RDP_HUMAN
20	40	72.7	614	1	RUI7_HUMAN
21	40	72.7	1106	1	STC_DROME
22	40	72.7	365	1	VG13_BPBO3
23	39	71.8	157	1	SSSP_PLAFO
24	39	70.9	813	1	DD13_HUMAN
25	39	70.9	171	1	MCRD_METBA
26	39	70.9	846	1	NRT_DROME
27	39	70.9	682	1	YB89_METJA
28	38	69.1	690	1	CNG1_BOVIN
29	38	69.1	609	1	FES_FSVGA
30	38	69.1	685	1	MOL1_CANAL
31	38	69.1	510	1	NT56_DROME
32	38	69.1	485	1	RBL_FLABI
33	38	69.1	485	1	RBL_FLAPR
34	38	69.1	485	1	RBL_HELAN
35	38	69.1	1644	1	RPA1_DROME
36	38	69.1	349	1	T2D6_HUMAN
37	38	69.1	972	1	TOP1_DROME
38	38	69.1	1249	1	TPP2_HUMAN
39	38	69.1	1262	1	TPP2_MOUSE
40	38	69.1	80	1	VPIP_BPT4
41	38	69.1	1045	1	Y052_HUMAN
42	38	69.1	198	1	Y530_METJA
43	38	69.1	1132	1	YK55_YEAST

ALIGNMENTS

```
RESULT 1
RUI7_XENLA
AC P09406;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE U1 SMALL NUCLEAR RIBONUCLEOPROTEIN 70 KD (U1 SNRNP 70 KD).
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89210819.
RA ETZERODT M., VIGNALI R., SCHERLY D., MATTAI I.W., CILIBERTO G.,
RA PHILIPSON L.;
RT *Structure and expression of a Xenopus gene encoding an snRNP protein
RT (U1 70K).";
RL EMBO J. 7:4311-4321(1988).
CC -1- FUNCTION: THIS PROTEIN IS ASSOCIATED WITH SN-RNP U1. IT BINDS
CC STEM LOOP I OF U1 SNRNA.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X12430; G65179; .
DR PIR; S02016; S02016.
DR PROSITE; PS00030; RNP_1; 1.
DR PFAM; PF00076; rrm; 1.
DR HSSP; P09651; 10P1.
KW NUCLEAR PROTEIN; RIBONUCLEOPROTEIN; RNA-BINDING.
FT DOMAIN 105 110 RNA-BINDING (RNP2) (BY SIMILARITY).
FT DOMAIN 147 154 RNA-BINDING (RNP1) (BY SIMILARITY).
FT DOMAIN 221 320 ARG/GLU-RICH (MIXED CHARGE).
FT DOMAIN 344 426 ARG/ASP/GLU-RICH (MIXED CHARGE).
SQ SEQUENCE 471 AA; 57203 MW; B97B3A81 CRC32;

Query Match 92.7%; Score 51; DB 1; Length 471;
Best Local Similarity 90.0%; Pred. No. 0.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKD 10
    |||||
Db 275 KDKDKDKD 284

RESULT 2
RDP_MOUSE
ID RDP_MOUSE STANDARD; PRT; 375 AA.
AC P19426;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE RD PROTEIN (WL623).
GN RD.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
```

RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE: 88178091.  
 RA LEVI-STRAUSS M., CARROLL M.C., STEINMETZ M., MEO T.;  
 RT "A previously undetected MHC gene with an unusual periodic  
 structure";  
 RL SCIENCE 240:201-204(1988).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- THE RD PROTEIN MIGHT NOT BIND RNA IN VIVO, POSSIBLY BECAUSE ITS  
 CC RNP CONSENSUS SEQUENCE IS HIGHLY ATYPICAL.  
 CC -!- THIS PROTEIN IS ENCODED IN THE MHC III LOCUS.  
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M21332; G199608; .  
 DR PIR; A40112; A40112.  
 DR MGD; MGI:102744; D17H6S45.  
 DR PFAM; PF00076; rtm; 1.  
 DR HSP; P19339; 1SXL.  
 KW RNA-BINDING; RIBONUCLEOPROTEIN; REPEAT; NUCLEAR PROTEIN;  
 KW MHC III.  
 FT DOMAIN 184 254 TANDEM REPEATS OF R-D.  
 FT DOMAIN 254 341 RNP DOMAIN.  
 SQ SEQUENCE 375 AA; 42555 MW; 660B0D44 CRC32;

Query Match 83.6%; Score 46; DB 1; Length 375;  
 Best Local Similarity 70.0%; Pred. No. 2.4;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 Db 212 RDRDRDKDKD 221

RESULT 3  
 FES\_FELCA STANDARD; PRT; 820 AA.  
 AC P14238;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (EC 2.7.1.112) (C-FES).  
 GN FES OR FPS.  
 OS FELIS SILVESTRI CATUS (CAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC CARNIVORA; FISSIPEDIA; FELIDAE; FELIS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 87198954.  
 RA ROEBROEK A.J.M., SCHALKEN J.A., ONNEKINK C., BLOEMERS H.P.J.,  
 RA VAN DE VEN W.J.M.;  
 RT "Structure of the feline c-fes/fps proto-oncogene: genesis of a  
 RT retroviral oncogene";  
 RL J. VIROL. 61:2009-2016(1987).  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M16705; G163848; .  
 DR EMBL; M16666; G163848; JOINED.  
 DR EMBL; M16667; G163848; JOINED.  
 DR EMBL; M16668; G163848; JOINED.  
 DR EMBL; M16669; G163848; JOINED.  
 DR EMBL; M16670; G163848; JOINED.  
 DR EMBL; M16671; G163848; JOINED.  
 DR EMBL; M16706; G163848; JOINED.  
 DR EMBL; M16672; G163848; JOINED.  
 DR EMBL; M16673; G163848; JOINED.  
 DR EMBL; M16674; G163848; JOINED.  
 DR EMBL; M16698; G163848; JOINED.  
 DR EMBL; M16700; G163848; JOINED.  
 DR EMBL; M16701; G163848; JOINED.  
 DR EMBL; M16702; G163848; JOINED.  
 DR EMBL; M16704; G163848; JOINED.  
 DR PIR; A27824; TVCTFF.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 DR PFAM; PF00017; SH2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00611; FCH; 1.  
 DR HSP; P11362; IFGI.  
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; PROTO-ONCOGENE; ATP-BINDING;  
 KW PHOSPHORYLATION; SH2 DOMAIN.  
 FT DOMAIN 458 547 SH2.  
 FT DOMAIN 559 820 PROTEIN KINASE.  
 FT NP\_BIND 565 573 ATP (BY SIMILARITY).  
 FT BINDING 588 588 ATP (BY SIMILARITY).  
 FT ACT\_SITE 681 681 BY SIMILARITY.  
 FT MOD\_RES 711 711 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 820 AA; 92974 MW; A9A030F8 CRC32;

Query Match 80.0%; Score 44; DB 1; Length 820;  
 Best Local Similarity 80.0%; Pred. No. 9.6;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 Db 159 KDKDRDKAKD 168

RESULT 4  
 FES\_HUMAN STANDARD; PRT; 822 AA.  
 AC P07332;  
 DT 01-APR-1988 (REL. 07, CREATED)  
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (EC 2.7.1.112) (C-FES).  
 GN FES OR FPS.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 90191711.  
 RA ALCALAY M., ANTONINI F., VAN DE VEN W.J., LANFRANCONE L.,  
 RA GRIGNANI F., PELICCI P.G.;  
 RT "Characterization of human and mouse c-fes cdna clones and  
 RT identification of the 5' end of the gene";  
 RL ONCOGENE 5:267-275(1990).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 86055727.  
 RA ROEBROEK A.J.M., SCHALKEN J.A., VERBEK J.S., VAN DEN OuwELAND A.M.W.,  
 RA ONNEKINK C., BLOEMERS H.P.J., VAN DE VEN W.J.M.;

RT "The structure of the human c-fes/fps proto-oncogene.";

RL EMBL J. 4:2897-2903(1985).  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X52192; G29891; -;  
 DR EMBL; X06292; G31349; -;  
 DR PIR; A24673; TVHUFF.  
 DR PIR; A60188; A60188.  
 DR MIN; 190030; -;  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PFAM; PF00017; SH2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00611; FCH; 1.  
 DR HSP; P11362; IFGI.  
 DR TRANSFERASE; TYROSINE-PROTEIN KINASE; PROTO-ONCOGENE; ATP-BINDING;  
 KW PHOSPHORYLATION; SH2 DOMAIN.  
 FT DOMAIN 460 549 SH2.  
 FT NP\_BIND 561 822 PROTEIN KINASE.  
 FT BINDING 567 575 ATP (BY SIMILARITY).  
 FT ACT\_SITE 590 590 ATP (BY SIMILARITY).  
 FT ACT\_SITE 683 683 BY SIMILARITY.  
 FT MOD\_RES 713 713 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CONFLICT 719 719 S -> L (IN REF. 2).  
 SQ SEQUENCE 822 AA; 93470 MW; 88669233 CRC32;

Query Match 80.08; Score 44; DB 1; Length 822;  
 Best Local Similarity 80.08; Pred. No. 9.6;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
 Db 161 KDKDKDKAKD 170  
 |||||:|:|

RESULT 5  
 FES\_MOUSE STANDARD; PRT; 820 AA.  
 AC P16879; Q2122;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (EC 2.7.1.112) (C-FES).  
 GN FES OR FPS.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89083198.  
 RA WILKS A.F., KURBAN R.R.;  
 RT "Isolation and structural analysis of murine c-fes cdna clones.";  
 RL ONCOGENE 3:289-294(1988).  
 RN [2]  
 RP SEQUENCE OF 678-747 FROM N.A.  
 RX MEDLINE; 90152381.  
 RA WILKS A.F., KURBAN R.R., HOVENS C.M., RALPH S.J.;  
 RT "The application of the polymerase chain reaction to cloning members

RT of the protein tyrosine kinase family.";  
 RL GENE 85:67-74(1989).  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN. BELONGS TO THE FES/FPS SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X12616; G50956; -;  
 DR EMBL; M33421; G200568; -;  
 DR MGD; MGI:95314; FES.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PFAM; PF00017; SH2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00611; FCH; 1.  
 DR HSP; P11362; IFGI.  
 DR TRANSFERASE; TYROSINE-PROTEIN KINASE; PROTO-ONCOGENE; ATP-BINDING;  
 KW PHOSPHORYLATION.  
 FT DOMAIN 458 547 SH2.  
 FT NP\_BIND 559 816 PROTEIN KINASE.  
 FT BINDING 565 573 ATP (BY SIMILARITY).  
 FT ACT\_SITE 588 588 ATP (BY SIMILARITY).  
 FT ACT\_SITE 681 681 BY SIMILARITY.  
 FT MOD\_RES 711 711 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CONFLICT 747 747 L -> P (IN REF. 2).  
 SQ SEQUENCE 820 AA; 93623 MW; EEA430F9 CRC32;

Query Match 80.08; Score 44; DB 1; Length 820;  
 Best Local Similarity 80.08; Pred. No. 9.6;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
 Db 161 KDKDKDKAKD 170  
 |||||:|:|

RESULT 6  
 PRSC\_HUMAN STANDARD; PRT; 324 AA.  
 AC P51665;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE 26S PROTEASOME REGULATORY SUBUNIT S12 (PROTEASOME SUBUNIT P40) (MOV34  
 DE PROTEIN).  
 GN PSMD7 OR MOV34L.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95275316.  
 RA TSURUMI C., DEMARTINO G.N., SLAUGHTER C., SHIMBARA N., TANAKA K.;  
 RT "cDNA cloning of p40, a regulatory subunit of the human 26S  
 RT proteasome, and a homolog of the Mov-34 gene product.";  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 210:600-608(1995).  
 CC -1- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME. WHICH  
 CC IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED  
 CC PROTEINS.  
 CC -1- SIMILARITY: TO OTHER SPECIES S12/MOV-34.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; D50063; G971270; -  
 DR MIM; 157970; -  
 KW PROTEASOME. 286 324 GLU/LYS-RICH (KEKE DOMAIN).  
 FT DOMAIN  
 SQ SEQUENCE 324 AA; 37060 MW; 63D97DBC CRC32;

Query Match 78.2%; Score 43; DB 1; Length 324;  
 Best Local Similarity 70.0%; Pred. No. 5.6;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKDCKDKDKD 10  
 Db 302 EDKDKDKDK 311

RESULT 7  
 ID T2FA\_DROME STANDARD; PRT; 577 AA.  
 AC Q05913;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT (TFIIF-ALPHA)  
 DE (TRANSCRIPTION FACTOR 5, LARGE CHAIN) (TF5A).  
 GN TFIIF-ALPHA.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93219133.  
 RA GONG D.-W., HORIKOSHI M., NAKATANI Y.;  
 RT "Analysis of cDNA encoding Drosophila transcription initiation factor  
 RT TFIIF alpha (RAP74).";  
 RL NUCLEIC ACIDS RES. 21:1492-1492(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93219101.  
 RA KEPHART D., PRICE M.P., BURTON Z.F., FINKELSTEIN A.,  
 RA GREENBLAU J., PRICE D.H.;  
 RT "Cloning of a Drosophila cDNA with sequence similarity to human  
 RT transcription factor RAP74".  
 RL NUCLEIC ACIDS RES. 21:1319-1319(1993).  
 RN [3]  
 RP FUNCTION.  
 RX MEDLINE; 94230464.  
 RA KEPHART D., WANG B.O., BURTON Z.F., PRICE D.H.;  
 RT "Functional analysis of Drosophila factor 5 (TFIIF), a general  
 RT transcription factor".  
 RL J. BIOL. CHEM. 269:13336-13543(1994).  
 CC -!- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT  
 CC BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE  
 CC INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES  
 CC TRANSCRIPTION ELONGATION.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L10331; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; X66982; G288191; -  
 DR PIR; S30237; S30237.  
 DR FLYBASE; FBgn0010282; TFIIF-alpha.  
 DR TRANSFAC; T02173; -  
 KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.  
 FT CONFLICT 20 20 R -> A (IN REF. 2).  
 FT CONFLICT 76 76 G -> E (IN REF. 2).  
 FT CONFLICT 528 528 A -> P (IN REF. 2).  
 SQ SEQUENCE 577 AA; 64453 MW; CA3DDSF6 CRC32;

Query Match 78.2%; Score 43; DB 1; Length 577;  
 Best Local Similarity 70.0%; Pred. No. 9.6;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKDCKDKDKD 10  
 Db 425 KDKDKDKDK 434

RESULT 8  
 ID NO60\_DROME STANDARD; PRT; 508 AA.  
 AC Q44081;  
 DT 15-DEC-1998 (REL. 37, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE NUCLEOLAR PROTEIN AT BAND 60B.  
 GN NOP60B.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA POOLE S.J., BILLIN A., CADWELL C., ERICKSON C., BUCHHOLZ R.,  
 RA MERRIAM J., CARBON J.;  
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.  
 CC -!- SIMILARITY: BELONGS TO THE TRUB FAMILY OF PSEUDOURIDINE SYNTHASES.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AF017230; G2661227; -  
 DR FLYBASE; FBgn0023184; Nop60B.  
 KW NUCLEAR PROTEIN.  
 FT DOMAIN 8 16 POLY-LYS.  
 FT DOMAIN 450 453 POLY-ALA.  
 FT DOMAIN 461 471 POLY-LYS.  
 FT DOMAIN 489 499 POLY-LYS.  
 SQ SEQUENCE 508 AA; 56830 MW; 09948431 CRC32;

Query Match 76.4%; Score 42; DB 1; Length 508;  
 Best Local Similarity 70.0%; Pred. No. 12;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKDCKDKDKD 10  
 Db 495 KKKDKDKDKD 504

RESULT 9

TOPI\_MOUSE  
 ID TOPI\_MOUSE STANDARD; PRT; 767 AA.  
 AC Q04750;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DNA TOPOISOMERASE I (EC 5.99.1.2).  
 GN TOPI OR TOP-1.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 RC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE; 93216125.  
 RA KOTWAI O., YASUI Y., SAKAI Y., WATANABE T., ISHII K.,  
 RA YANAGIHARA S., ANDOH T.;  
 RT "Cloning of the mouse cDNA encoding DNA topoisomerase I and  
 RT chromosomal location of the gene";  
 RL GENE 125:211-216(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA HUI C.F., LO C.K., HWANG J.;  
 RL SUBMITTED (JUL-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE  
 CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.  
 CC -!- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED  
 CC DNA, FOLLOWED BY PASSAGE AND REJOINING.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH NEGATIVE AND  
 CC POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES RELAX ONLY  
 CC NEGATIVE SUPERCOILS.  
 CC -!- WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT  
 CC SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN  
 CC IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE  
 CC ENZYME-SEVERED DNA STRAND.  
 CC -!- SIMILARITY: BELONGS TO EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; D10061; D1001416;  
 DR EMBL; L20632; G309523;  
 DR PIR; J00144; J00144.  
 DR MGD; MGI:98788; TOPI.  
 DR PFAM; PF01028; TOPOISOMERASE\_I\_EUK; 1.  
 DR HSSP; P11387; 1A35.  
 KW ISOMERASE; TOPOISOMERASE; DNA-BINDING.  
 FT ACT\_SITE 725 725 DNA CLEAVAGE (BY SIMILARITY).  
 FT CONFLICT 91 91 R -> P (IN REF. 2).  
 FT CONFLICT 121 121 R -> E (IN REF. 2).  
 FT CONFLICT 129 129 A -> V (IN REF. 2).  
 FT CONFLICT 161 161 MISSING (IN REF. 2).  
 FT CONFLICT 167 167 S -> L (IN REF. 2).  
 FT CONFLICT 277 277 R -> W (IN REF. 2).  
 FT CONFLICT 292 292 G -> E (IN REF. 2).  
 FT CONFLICT 522 522 G -> V (IN REF. 2).  
 FT CONFLICT 533 533 G -> W (IN REF. 2).  
 FT CONFLICT 762 762 D -> Y (IN REF. 2).  
 SQ SEQUENCE 767 AA; 90789 MW; 560FCA80 CRC32;

Query Match 76.4%; Score 42; DB 1; Length 767;  
 Best Local Similarity 77.8%; Pred. No. 17;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDK 9  
 |||||||:|

Db 40 KDKDKREK 48  
 RESULT 10  
 CRT1\_BOVIN  
 ID CRT1\_BOVIN STANDARD; PRT; 400 AA.  
 AC P52193;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE CALRETICULIN, BRAIN ISOFORM 1 (CRP55) (CALREGULIN) (HACBP).  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 94183174.  
 RA MATSUOKA K., SETA K., YAMAKAWA Y., OKUYAMA T., SHINODA T., ISOBE T.;  
 RT "Covalent structure of bovine brain calreticulin";  
 RL BIOCHEM. J. 298:435-442(1994).  
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND  
 CC LOW AFFINITY CALCIUM-BINDING SITES.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.  
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.  
 CC PROSITE; PS0014; ER\_TARGET; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PFAM; PF00262; calreticulin; 1.  
 DR ENDOPLASMIC RETICULUM; CALCIUM-BINDING; REPEAT; GLYCOPROTEIN.  
 FT DOMAIN 1 180 N-DOMAIN.  
 FT DOMAIN 181 291 P-DOMAIN.  
 FT DOMAIN 292 400 C-DOMAIN.  
 FT DOMAIN 174 238 4 X APPROXIMATE REPEATS.  
 FT REPEAT 174 185 1-1.  
 FT REPEAT 193 204 1-2.  
 FT REPEAT 210 221 1-3.  
 FT REPEAT 227 238 1-4.  
 FT DOMAIN 242 280 3 X APPROXIMATE REPEATS.  
 FT REPEAT 242 252 2-1.  
 FT REPEAT 256 266 2-2.  
 FT REPEAT 270 280 2-3.  
 FT DOMAIN 334 390 ASP/GLU/LYS-RICH.  
 FT DISULFID 120 146  
 FT CARBOHYD 162 162  
 FT SITE 397 400  
 SQ SEQUENCE 400 AA; 46381 MW; 023936E0 CRC32;

Query Match 74.5%; Score 41; DB 1; Length 400;  
 Best Local Similarity 70.0%; Pred. No. 13;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKDK 10  
 |||:|:|:|

Db 367 KDDDEDKDED 376

RESULT 11  
 CRT2\_BOVIN  
 ID CRT2\_BOVIN STANDARD; PRT; 421 AA.  
 AC P42918;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CALRETICULIN, BRAIN ISOFORM 2 PRECURSOR (CRP55) (CALREGULIN) (HACBP).  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;  
 RX MEDLINE; 93385184.  
 RA LIU N., FINE R.E., JOHNSON R.J.:  
 RT "Comparison of cDNAs from bovine brain coding for two isoforms of  
 RT calreticulin";  
 RL BIOCHIM. BIOPHYS. ACTA 1202:70-76(1993).  
 CC -!- FUNCTION: THIS PROTEIN BINDS CALCIUM. THERE ARE BOTH HIGH AND  
 CC LOW AFFINITY CALCIUM-BINDING SITES.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.  
 CC -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L13462; G348694;  
 DR PROSITE; PS00014; ER-TARGET; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PFAM; PF00262; calreticulin; 1.  
 KW ENDOPLASMIC RETICULUM; CALCIUM-BINDING; REPEAT; SIGNAL.  
 FT SIGNAL 1 34  
 FT CHAIN 35 421 CALRETICULIN, BRAIN ISOFORM 2.  
 FT DOMAIN 35 201 N-DOMAIN.  
 FT DOMAIN 202 312 P-DOMAIN.  
 FT DOMAIN 313 421 C-DOMAIN.  
 FT DOMAIN 195 259 4 X APPROXIMATE REPEATS.  
 FT REPEAT 195 206 1-1.  
 FT REPEAT 214 225 1-2.  
 FT REPEAT 231 242 1-3.  
 FT REPEAT 248 259 1-4.  
 FT DOMAIN 263 301 3 X APPROXIMATE REPEATS.  
 FT REPEAT 263 273 2-1.  
 FT REPEAT 277 287 2-2.  
 FT REPEAT 291 301 2-3.  
 FT DOMAIN 366 411 ASP/GLU/LYS-RICH.  
 FT DISULFID 141 167 BY SIMILARITY.  
 FT CARBOHYD 183 183 POTENTIAL.  
 FT SITE 418 421 PREVENT SECRETION FROM ER.  
 SQ SEQUENCE 421 AA; 48812 MW; 78418488 CRC32;  
  
 Query Match 74.5%; Score 41; DB 1; Length 421;  
 Best Local Similarity 70.0%; Pred. No. 14;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 KDKDKDKDKD 10  
 Db 388 KDDDEKDEKDED 397  
  
 RESULT 12  
 FSYA\_DROME STANDARD; PRT; 696 AA.  
 AC P25028;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE MITOSIS INITIATION PROTEIN FS(1)YA.  
 GN FS(1)YA.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91098653.

RA LIN H., WOLFNER M.F.:  
 RT "The Drosophila maternal-effect gene fs(1)Ya encodes a cell cycle-  
 RT dependent nuclear envelope component required for embryonic  
 RT mitosis.";  
 RL CELL 64:49-62(1991).  
 RN [2]  
 RP REVISIONS.  
 RA LIU J., WOLFNER M.F.:  
 RL SUBMITTED (JAN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: CELL CYCLE-DEPENDENT NUCLEAR ENVELOPE COMPONENT  
 CC REQUIRED FOR EMBRYONIC MITOSIS.  
 CC -!- SUBCELLULAR LOCATION: IN THE NUCLEAR ENVELOPE DURING INTERPHASE TO  
 CC METAPHASE, AND IN THE NUCLEOLUS AND CYTOPLASM DURING ANAPHASE  
 CC AND TELOPHASE.  
 CC -!- DEVELOPMENTAL STAGE: MATERNAL PROTEIN SYNTHESIZED DURING  
 CC POSTGENETIC MATURATION AND PERSISTING THROUGHOUT EMBRYOGENESIS.  
 CC -!- THE OPA REPEAT-CONTAINING AND THE SER/THR-RICH REGIONS MIGHT BE  
 CC INVOLVED IN PROTEIN-PROTEIN INTERACTIONS WITH THE MAJOR PROTEIN-  
 CC RICH LAYER OF THE NUCLEAR ENVELOPE.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M38442; G552093;  
 DR PIR; A38436; A38436.  
 DR FLYBASE; FBgn0000927; fs(1)Ya.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; UNKNOWN 1.  
 KW CELL CYCLE; MITOSIS; NUCLEAR PROTEIN; PHOSPHORYLATION.  
 FT DOMAIN 270 281 POLY-GLN (OPA-REPEAT).  
 FT DOMAIN 512 520 NUCLEAR LOCALIZATION SIGNAL (PROBABLE).  
 FT DOMAIN 534 538 NUCLEAR LOCALIZATION SIGNAL (PROBABLE).  
 FT DOMAIN 448 696 RICH IN CHARGED AA.  
 SQ SEQUENCE 696 AA; 77709 MW; 1A1631DF CRC32;  
  
 Query Match 74.5%; Score 41; DB 1; Length 696;  
 Best Local Similarity 60.0%; Pred. No. 22;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 KDKDKDKDKD 10  
 Db 612 KDRDRDRSD 621  
  
 RESULT 13  
 SPCO\_HUMAN STANDARD; PRT; 2364 AA.  
 AC Q01082; Q16057;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE SPECTRIN BETA CHAIN, BRAIN (SPECTRIN, NON-ERYTHROID BETA CHAIN)  
 DE (FODRIN BETA CHAIN) (SPTBN1).  
 GN SPTB2.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 92406787.  
 RA HU R.J., WATANABE M., BENNETT V.:  
 RT "Characterization of human brain cDNA encoding the general isoform of  
 RT beta-spectrin.";  
 RL J. BIOL. CHEM. 267:18715-18722(1992).  
 RN [2]  
 RP SEQUENCE OF 293-1544 FROM N.A.  
 RX MEDLINE; 94010920.



RA CHANG J.G., SCARPA A., EDDY R.L., BYERS M.G., HARRIS A.S.,  
 RA MORROW J.S., WATKINS P., SHOWS T.B., FORGET B.G.;  
 RT "Cloning of a portion of the chromosomal gene and cDNA for human  
 RT beta-fodrin, the nonerythroid form of beta-spectrin.";  
 RL GENOMICS 17:287-293(1993).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 173-280.  
 RX MEDLINE: 97307247.  
 RA CARUGO K.D., BANUELOS S., SARASTE M.;  
 RT "Crystal structure of a calponin homology domain.";  
 RL NAT. STRUCT. BIOL. 4:175-179(1997).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS) OF 173-281.  
 RA BANUELOS S., SARASTE M., CARUGO K.D.;  
 RL SUBMITTED (JUL-1998) TO THE PDB DATA BANK.  
 CC -!- FUNCTION: FODRIN, WHICH SEEMS TO BE INVOLVED IN SECRETION,  
 CC INTERACTS WITH CALMODULIN IN A CALCIUM-DEPENDENT MANNER AND IS  
 CC THUS CANDIDATE FOR THE CALCIUM-DEPENDENT MOVEMENT OF THE  
 CC CYTOSKELETON AT THE MEMBRANE.  
 CC SUBUNIT: LIKE ERYTHROCYTE SPECTRIN, THE SPECTRIN-LIKE PROTEINS  
 CC ARE CAPABLE TO FORM DIMERS WHICH CAN FURTHER ASSOCIATE TO  
 CC TETRAMERS.  
 CC -!- DOMAIN: SPECTRIN-LIKE PROTEINS HAVE FIVE DOMAINS:  
 CC (1) N-TERMINAL DOMAIN (N),  
 CC (2) DOMAIN BETWEEN THE N-TERMINAL AND MIDDLE DOMAIN (NM),  
 CC (3) MIDDLE DOMAIN (M),  
 CC (4) DOMAIN BETWEEN THE MIDDLE AND C-TERMINAL DOMAIN (MC),  
 CC (5) C-TERMINAL DOMAIN (C).  
 CC NM AND MC DOMAINS ARE COMPOSED OF TYPICAL SPECTRIN 106 RESIDUES  
 CC REPEATS (1-8 FOR NM AND 12-19 FOR MC) AND ARE HOMOLOGOUS TO EACH  
 CC OTHER. N, M, AND C DOMAINS ARE COMPOSED OF SEQUENCES THAT DO NOT  
 CC FORM TYPICAL SPECTRIN REPEATS.  
 CC -!- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M96803; G338443; -.  
 DR EMBL: S65762; G425553; -.  
 DR PDB: 1AA2; 04-FEB-98.  
 DR PDB: 1BKR; 15-JUL-98.  
 DR MM; 182790; -.  
 DR PROSITE: PS00019; ACTININ\_1; 1.  
 DR PROSITE: PS00020; ACTININ\_2; 1.  
 DR PROSITE: PS50003; PH\_DOMAIN; 1.  
 DR PFAM: PF00169; PH; 1.  
 DR PFAM: PF00307; actinin-binding; 1.  
 DR PFAM: PF00435; spectrin; 17.  
 DR CYTOSKELETON; MEMBRANE; REPEAT: ACTIN-BINDING; CAPPING PROTEIN;  
 KW 3D-STRUCTURE.  
 FT DOMAIN 1 272 ACTIN-BINDING (BY SIMILARITY).  
 FT REPEAT 273 384 1.  
 FT REPEAT 385 498 2.  
 FT REPEAT 499 608 3.  
 FT REPEAT 609 714 4.  
 FT REPEAT 715 819 5.  
 FT REPEAT 820 925 6.  
 FT REPEAT 926 1032 7.  
 FT REPEAT 1033 1139 8.  
 FT REPEAT 1140 1245 9.  
 FT REPEAT 1246 1350 10.  
 FT REPEAT 1351 1462 11.  
 FT REPEAT 1463 1562 12.  
 FT REPEAT 1563 1668 13.  
 FT REPEAT 1669 1775 14.  
 FT REPEAT 1776 1881 15.  
 FT REPEAT 1882 1987 16.

FT REPEAT 1988 2133 17.  
 FT DOMAIN 2137 2307 PH.  
 FT CONFLICT 1411 1411 H -> D (IN REF. 2).  
 SQ SEQUENCE 2364 AA; 274629 MW; D45B9267 CRC32;  
 Query Match 74.5%; Score 41; DB 1; Length 2364;  
 Best Local Similarity 77.8%; Pred. No. 67;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KKDKDKDK 9  
 Db 2347 KKDKDKDK 2355  
 RESULT 14  
 ID TOP1\_CRIGR STANDARD; PRT; 767 AA.  
 AC Q07050;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DNA TOPOISOMERASE I (EC 5.99.1.2).  
 GN TOPI OR TOP-1.  
 OS CRICETULUS GRISEUS (CHINESE HAMSTER).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; CRICETINAE; CRICETULUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94064611.  
 RA TANIZAWA A., BERTRAND R., KOHLHAGEN G., TABUCHI A., JENKINS J.,  
 RA POMIER Y.;  
 RT "Cloning of Chinese hamster DNA topoisomerase I cDNA and  
 RT identification of a single point mutation responsible for  
 RT camptothecin resistance.";  
 RL J. BIOL. CHEM. 268:25463-25468(1993).  
 CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE  
 CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.  
 CC -!- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED  
 CC DNA, FOLLOWED BY PASSAGE AND REJOINING.  
 CC -!- ENZYME REGULATION: SPECIFICALLY INHIBITED BY CAMPTOTHECIN (CPT), A  
 CC PLANT ALKALOID WITH ANTITUMOR ACTIVITY.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH NEGATIVE AND  
 CC POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES RELAX ONLY  
 CC NEGATIVE SUPERCOILS.  
 CC -!- WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT  
 CC SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN  
 CC IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE  
 CC ENZYME-SEVERED DNA STRAND.  
 CC -!- SIMILARITY: BELONGS TO EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 221624; G297077; -.  
 DR EMBL: 221625; G297079; -.  
 DR PIR: S32697; S32697.  
 DR PROSITE: PS00176; TOPOISOMERASE\_I\_EUK; 1.  
 DR PFAM: PF01028; Topoisomerase\_I; 1.  
 DR HSSP: P11387; 1A35.  
 KW ISOMERASE; TOPOISOMERASE; DNA-BINDING.  
 FT ACT\_SITE 725 725 DNA CLEAVAGE (BY SIMILARITY).  
 FT VARIANT 505 505 S -> G (IN CPT-RESISTANT CELL).  
 SQ SEQUENCE 767 AA; 90867 MW; 333EDD67 CRC32;  
 Query Match 74.5%; Score 41; DB 1; Length 767;

Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10  
III IIII:  
Db 80 KDKHKDRDKE 89

## RESULT 15

TOP1\_HUMAN STANDARD; PRT; 765 AA.  
AC P11387;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA TOPOISOMERASE I (EC 5.99.1.2).  
GN TOP1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
CC [1]  
CC SEQUENCE FROM N.A.  
RX MEDLINE; 88190108.  
RA D'ARPA P., MACHLIN P.S., PATRIE H. III, ROTHFIELD N.F.,  
RA CLEVELAND D.W., EARNSHAW W.C.;  
RT "cDNA cloning of human DNA topoisomerase I: catalytic activity of a  
67.7-kDa carboxyl-terminal fragment.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 85:2543-2547(1988).  
CC [2]  
CC SEQUENCE FROM N.A.  
RX MEDLINE; 91236733.  
RA KUNZE N., YANG G., DOLBERG M., SUNDARP R., KNIPPERS R., RICHTER A.;  
RT "Structure of the human type I DNA topoisomerase gene.";  
RL J. BIOL. CHEM. 266:9610-9616(1991).  
CC [3]  
CC SEQUENCE OF 541-765 FROM N.A.  
RX MEDLINE; 89288043.  
RA ZHOU B.S., BASTOW K.F., CHENG Y.C.;  
RT "Characterization of the 3' region of the human DNA topoisomerase I  
gene.";  
RL CANCER RES. 49:3922-3927(1989).  
CC [4]  
CC SEQUENCE OF 657-765 FROM N.A.  
RX MEDLINE; 90046823.  
RA MAUL G.G., JIMENEZ S.A., RIGGS E., ZIEMNICKA-KOTULA D.;  
RT "Determination of an epitope of the diffuse systemic sclerosis marker  
antigen DNA topoisomerase I: sequence similarity with retroviral  
p30gag protein suggests a possible cause for autoimmunity in systemic  
sclerosis.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:8492-8496(1989).  
CC [5]  
CC VARIANTS CPT-RESISTANT.  
RX MEDLINE; 91187651.  
RA TAMURA H., KOHCHI C., YAMADA R., IKEDA T., KOIWA O., PATTERSON E.,  
RA KEENE J.D., OKADA K., KJELSEN E., NISHIKAWA K.;  
RT "Molecular cloning of a cDNA of a camptothecin-resistant human DNA  
topoisomerase I and identification of mutation sites.";  
RL NUCLEIC ACIDS RES. 19:69-75(1991).  
CC [6]  
CC VARIANT CPT-RESISTANT ALA-729.  
RX MEDLINE; 93075133.  
RA KUBOTA N., KANZAWA F., NISHIO K., TAKEDA Y., OHMORI T., FUJIWARA Y.,  
RA TERASHIMA Y., SAIGO N.;  
RT "Detection of topoisomerase I gene point mutation in CPT-11 resistant  
lung cancer cell line.";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 188:571-577(1992).  
CC [7]  
CC X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 215-765.  
RX MEDLINE; 98155246.  
RA REDINBO M.R., STEWART L., KUHN P., CHAMPOUX J.J., HOL W.G.J.;  
RT "Crystal structures of human topoisomerase I in covalent and  
noncovalent complexes with DNA.";  
RL SCIENCE 279:1504-1513(1998).

[8]  
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 215-765.  
RX MEDLINE; 98155254.  
RA STEWART L., REDINBO M.R., QIU X., HOL W.G.J., CHAMPOUX J.J.;  
RT "A model for the mechanism of human topoisomerase I.";  
RL SCIENCE 279:1534-1541(1998).  
CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE  
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.  
CC -!- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED  
CC DNA, FOLLOWED BY PASSAGE AND REJOINING.  
CC -!- ENZYME REGULATION: SPECIFICALLY INHIBITED BY CAMPTOTHECIN (CPT), A  
CC PLANT ALKALOID WITH ANTITUMOR ACTIVITY.  
CC -!- SUBUNIT: MONOMER.  
CC -!- EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH NEGATIVE AND  
CC POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES RELAX ONLY  
CC NEGATIVE SUPERCOILS.  
CC -!- WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT  
CC SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN  
CC IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE  
CC ENZYME-SEVERED DNA STRAND.  
CC -!- SIMILARITY: BELONGS TO EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J03250; G339806; -;  
DR EMBL; M60706; G339804; -;  
DR EMBL; M60688; G339804; JOINED.  
DR EMBL; M60689; G339804; JOINED.  
DR EMBL; M60690; G339804; JOINED.  
DR EMBL; M60691; G339804; JOINED.  
DR EMBL; M60692; G339804; JOINED.  
DR EMBL; M60693; G339804; JOINED.  
DR EMBL; M60694; G339804; JOINED.  
DR EMBL; M60695; G339804; JOINED.  
DR EMBL; M60696; G339804; JOINED.  
DR EMBL; M60697; G339804; JOINED.  
DR EMBL; M60698; G339804; JOINED.  
DR EMBL; M60699; G339804; JOINED.  
DR EMBL; M60700; G339804; JOINED.  
DR EMBL; M60701; G339804; JOINED.  
DR EMBL; M60702; G339804; JOINED.  
DR EMBL; M60703; G339804; JOINED.  
DR EMBL; M60704; G339804; JOINED.  
DR EMBL; M60705; G339804; JOINED.  
DR EMBL; X16479; -; NOT\_ANNOTATED\_CDS.  
DR PIR; M27913; G339808; -;  
DR PIR; A30887; ISHUT1.  
DR PDB; 1A31; 19-AUG-98.  
DR PDB; 1A35; 26-AUG-98.  
DR PDB; 1A36; 12-AUG-98.  
DR AARHUS/GHENT-2DPAGE; 610; NEPHGE.  
DR MIM; 126420; -;  
DR PROSITE; PS00176; TOPOISOMERASE\_I\_EUK; 1.  
DR PFAM; PF01028; Topoisomerase\_I\_1  
KW ISOMERASE; TOPOISOMERASE; DNA-BINDING; POLYMORPHISM; 3D-STRUCTURE.  
FT ACT\_SITE 723 723 DNA CLEAVAGE (BY SIMILARITY).  
FT VARIANT 533 533 D -> G (IN CPT-RESISTANT CELL).  
FT VARIANT 729 729 T -> A (IN CPT-RESISTANT CELL).  
FT CONFLICT 145 145 V -> A (IN REF. 2).  
SQ SEQUENCE 765 AA; 90753 MW; ED69D4F3 CRC32;

Query Match 74.5%; Score 41; DB 1; Length 765;  
Best Local Similarity 70.0%; Pred. No. 24;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10

Db        III III:II:  
          78 KDKHKORDKE 87

Search completed: September 7, 1999, 23:50:04  
Job time: 1943 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:30 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-4  
Perfect score: 55  
Sequence: 1 KDKDKDKDKD 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL\_10:  
1: sp-archaea:  
2: sp-bacteria:  
3: sp-fungi:  
4: sp-human:  
5: sp-invertebrate:  
6: sp-mammal:  
7: sp-mhc:  
8: sp-organelle:  
9: sp-phage:  
10: sp-plant:  
11: sp-rodent:  
12: sp-virus:  
13: sp-vertebrate:  
14: sp-unclassified:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	92.7	1997	5	O96239	O96239 plasmodium
2	51	92.7	257	10	P93751	P93751 arabidopsis
3	51	92.7	239	10	Q39267	Q39267 arabidopsis
4	50	90.9	160	5	Q26108	Q26108 plasmodium
5	46	83.6	1566	4	O43810	O43810 homo sapien
6	46	83.6	1581	4	O75447	O75447 homo sapien
7	46	83.6	2559	5	O44113	O44113 drosophila
8	46	83.6	2559	5	O44381	O44381 drosophila
9	46	83.6	44	11	P97485	P97485 mus musculus
10	46	83.6	50	11	P97486	P97486 mus musculus
11	46	83.6	152	13	O9Y128	O9Y128 carassius a
12	45	81.8	1034	5	O16015	O16015 drosophila
13	45	81.8	1050	5	O09655	O09655 caenorhabdi
14	45	81.8	2206	5	O96205	O96205 plasmodium
15	45	81.8	737	10	O80716	O80716 arabidopsis
16	45	81.8	281	12	Q84947	Q84947 papaya ring
17	45	81.8	345	12	Q89336	Q89336 zucchini ye
18	45	81.8	326	13	O91648	O91648 xenopus lae
19	44	80.0	22	5	O04123	O04123 plasmodium
20	44	80.0	1257	5	O22944	O22944 caenorhabdi
21	44	80.0	766	13	P79994	P79994 gallus gall
22	43	78.2	644	2	Q48034	Q48034 haemophilus
23	43	78.2	660	2	Q57443	Q57443 haemophilus
24	43	78.2	776	5	O15793	O15793 p strain dd
25	43	78.2	597	5	O20108	O20108 caenorhabdi
26	43	78.2	3498	5	O20497	O20497 caenorhabdi
27	43	78.2	2911	5	O93442	O93442 caenorhabdi
28	43	78.2	480	8	P94079	P94079 psilanthus
29	43	78.2	733	10	P93008	P93008 arabidopsis

ALIGNMENTS

RESULT 1

O96239  
ID O96239 PRELIMINARY; PRT: 1997 AA.  
AC O96239;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE BRAHMA ORTHOLOG (DNA HELICASE SUPERFAMILY II).  
GN PFS0730W.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOO S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTERSON M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
RT falciparum."  
RL Science 282:1126-1132(1998).  
DR EMBL; AE001414; AAC71935.1; -.  
KW Helicase.  
SQ SEQUENCE 1997 AA; 237125 MW; 673C5FCB CRC32;

Query Match 92.7%; Score 51; DB 5; Length 1997;  
Best Local Similarity 90.0%; Pred. No. 3.1;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KDKDKDKDKD 10

Db 1618 KDKDKDKDKD 1627

RESULT 2

P93751  
ID P93751 PRELIMINARY; PRT: 257 AA.  
AC P93751;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE ZINC FINGER PROTEIN 8, ZFP8.  
GN T06D20.16.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eukaryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;

RA ROUNSLEY S.D., LIN X., KETCHUM K.A., PHILLIPS C.A., BRANDON R.C.,  
 RA FUHRMANN J.L., WHITE O., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R.,  
 RA VENTER J.C.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U90439; AAB63548.1; -;  
 DR PFAM: PF00096; zf-C2H2.1;  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 1;  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 257 AA; 29571 MW; FB769546 CRC32;

Query Match 92.7%; Score 51; DB 10; Length 257;  
 Best Local Similarity 90.0%; Pred. No. 0.5;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 DB 75 KDKDKDKDKD 84

RESULT 3  
 Q39267 PRELIMINARY; PRT; 239 AA.  
 AC Q39267;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE ZINC FINGER PROTEIN (FRAGMENT).  
 GN ZFP8.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LANDSBERG RECTA; TISSUE-LEAF;  
 RX MEDLINE; 95322589.  
 RA TAGUE B.W., GOODMAN H.M.;  
 RT "Characterization of a family of Arabidopsis zinc finger protein  
 cDNAs.";  
 RL Plant Mol. Biol. 28:267-279(1995).  
 DR EMBL: L39651; AAA87304.1; -;  
 DR PFAM: PF00096; zf-C2H2.1;  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 1.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 FT NON\_TER 1  
 SQ SEQUENCE 239 AA; 27781 MW; 48976A9F CRC32;

Query Match 92.7%; Score 51; DB 10; Length 239;  
 Best Local Similarity 90.0%; Pred. No. 0.47;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 DB 73 KDKDKDKDKD 82

RESULT 4  
 Q26108 PRELIMINARY; PRT; 160 AA.  
 AC Q26108;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE SPOROZOITE ANTIGEN PRECURSOR.  
 OS Plasmodium reichenowi.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MOELANS I., LAL A.A., KONINGS R.N., SCHOENMAKERS J.G.;  
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

RN SEQUENCE FROM N.A.  
 RP MEDLINE; 92158015.  
 RA MOELANS I., LAL A.A., KONINGS R.N., SCHOENMAKERS J.G.;  
 RT "Sequence of a 16-kilodalton sexual stage and sporozoite surface  
 antigen of Plasmodium reichenowi and comparison with Pf16 of  
 Plasmodium falciparum.";  
 RL Mol. Biochem. Parasitol. 50:349-350(1992).  
 DR EMBL: M75444; AAA29766.1; -;  
 KW Signal; Sporozoite.  
 FT SIGNAL 22  
 FT CHAIN 23 160 SPOROZOITE ANTIGEN.  
 SQ SEQUENCE 160 AA; 17025 MW; 863EFD89 CRC32;

Query Match 90.9%; Score 50; DB 5; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 0.46;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKDKDKDKD 10  
 DB 141 DKDKDKDKD 149

RESULT 5  
 Q43810 PRELIMINARY; PRT; 1566 AA.  
 AC Q43810;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE RB18A PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-HEART;  
 RX MEDLINE; 98105695.  
 RA DRANE P., BAREL M., BALBO M., FRADE R.;  
 RT "Identification of RB18A, a 205 kDa new p53 regulatory protein which  
 shares antigenic and functional properties with p53.";  
 RL Oncogene 15:3013-3024(1997).  
 DR EMBL: Y13467; CAA73867.1; -;  
 SQ SEQUENCE 1566 AA; 166568 MW; 9EBA89E9 CRC32;

Query Match 83.6%; Score 46; DB 4; Length 1566;  
 Best Local Similarity 70.0%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
 DB 1493 KDKDKDKDKD 1502

RESULT 6  
 Q75447 PRELIMINARY; PRT; 1581 AA.  
 AC Q75447;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE THYROID HORMONE RECEPTOR-ASSOCIATED PROTEIN COMPLEX COMPONENT  
 DE TRAP220.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA YUAN C.X., ITO M., FONDELL J.D., FU Z.Y., ROEDER R.G.;  
 RL "The TRAP220 component of a thyroid hormone receptor- associated

RT protein (TRAP) coactivator complex interacts directly with nuclear  
RL receptors in a ligand-dependent fashion.;  
RL Proc. Natl. Acad. Sci. U.S.A. 95:7939-7944(1998).  
DR EMBL; AF055994; AAC39854.1; -.  
SQ SEQUENCE 1581 AA; 168517 MW; A0DA5415 CRC32;

Query Match 83.6%; Score 46; DB 4; Length 1581;  
Best Local Similarity 70.0%; Pred. No. 14;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
Db 1508 KDKDKDKDKD 1517  
:|:|:|:|:|:|

RESULT 7  
O44113 PRELIMINARY; PRT; 2559 AA.  
AC O44113;

DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE PUTATIVE GUANINE NUCLEOTIDE EXCHANGE FACTOR RHOGEF2.  
GN RHOGEF2.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
OC Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.

RA HAECKER U., PERRIMON N.;  
RL Genes Dev. 0:0-0(1998).  
DR EMBL; AF031930; AAB88816.1; -.  
DR FLYBASE; FBgn0023172; RHOGEF2.  
DR PFAM; PF00130; DAG\_PE-bind; 1.  
DR PFAM; PF00595; PDZ; 1.  
DR PFAM; PF00621; RHOGEF; 1.  
DR PROSITE; PS00479; DAG\_PE\_BINDING\_DOMAIN; 1.  
KW Phorbol-ester binding.  
SQ SEQUENCE 2559 AA; 280946 MW; 275758B8 CRC32;

Query Match 83.6%; Score 46; DB 5; Length 2559;  
Best Local Similarity 70.0%; Pred. No. 21;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
Db 130 RDRDKDKDKD 139  
:|:|:|:|:|:|

RESULT 8  
O44381 PRELIMINARY; PRT; 2559 AA.  
AC O44381;

DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE SHAR PEI/DHOGEF2.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
OC Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE; 98088790.  
RA BARRETT K., LEPTIN M., SETTLEMAN J.;  
RT "The Rho GTPase and a putative RHOGEF mediate a signaling pathway for  
the cell shape changes in Drosophila gastrulation.";  
RL Cell 91:905-915(1997).  
DR EMBL; AF032870; AAC38820.1; -.  
DR FLYBASE; FBgn0023172; RHOGEF2.

DR PFAM; PF00130; DAG\_PE-bind; 1.  
DR PFAM; PF00595; PDZ; 1.  
DR PFAM; PF00621; RHOGEF; 1.  
DR PROSITE; PS00479; DAG\_PE\_BINDING\_DOMAIN; 1.  
KW Phorbol-ester binding.  
SQ SEQUENCE 2559 AA; 280857 MW; 3A2AC8F9 CRC32;

Query Match 83.6%; Score 46; DB 5; Length 2559;  
Best Local Similarity 70.0%; Pred. No. 21;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
Db 130 RDRDKDKDKD 139  
:|:|:|:|:|:|

RESULT 9  
P97485 PRELIMINARY; PRT; 44 AA.  
AC P97485;

DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-MAY-1997 (TREMBlrel. 03, Last annotation update)  
DE H-2B HAPLOTYPE RD (FRAGMENT).  
GN RD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-129/SV;  
RA TAYLOR P.R., SLINGSBY J.H., WALPORT M.J., BOTTO M.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U83841; AAB41535.1; -.  
FT NON\_TER 1  
FT NON\_TER 44  
SQ SEQUENCE 44 AA; 5944 MW; 4811E804 CRC32;

Query Match 83.6%; Score 46; DB 11; Length 44;  
Best Local Similarity 70.0%; Pred. No. 0.55;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDKDKDKDKD 10  
Db 17 RDRDKDKDKD 26  
:|:|:|:|:|:|

RESULT 10  
P97486 PRELIMINARY; PRT; 50 AA.  
AC P97486;

DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-MAY-1997 (TREMBlrel. 03, Last annotation update)  
DE H-2R HAPLOTYPE RD (FRAGMENT).  
GN RD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LP.R.III/SN;

RA TAYLOR P.R., SLINGSBY J.H., WALPORT M.J., BOTTO M.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U83842; AAB41536.1; -.  
FT NON\_TER 1  
FT NON\_TER 50  
SQ SEQUENCE 50 AA; 6744 MW; 06837E29 CRC32;

Query Match 83.6%; Score 46; DB 11; Length 50;

Best Local Similarity 70.0%; Pred. No. 0.62; Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
      :|:|:|:|:|  
Db 15 RDRDRDKDKD 24

RESULT 11  
QYI28 PRELIMINARY; PRT; 152 AA.  
AC QYI28;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE PUTATIVE ARGININE-RICH NUCLEAR PROTEIN (FRAGMENT).  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii;  
OC Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae;  
OC Cyprinidae; Cyprininae; Carassius.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PITUITARY;  
RX MEDLINE; 99049263.  
RA BLAZQUEZ M., BOSMA P.T., CHANG J.P., DOCHERTY K., TRUDEAU V.L.;  
RT "Gamma-aminobutyric acid up-regulates the expression of a novel  
RT secretogranin-II messenger ribonucleic acid in the goldfish  
RT pituitary.";  
RL Endocrinology 139:4870-4880(1998).  
DR EMBL; AF046006; AAC94998.1; -.  
KW Nuclear protein.  
FT NON\_TER 1  
SQ SEQUENCE 152 AA; 18050 MW; 01AFFA5B CRC32;

Query Match 83.6%; Score 46; DB 13; Length 152;  
Best Local Similarity 70.0%; Pred. No. 1.7;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
      :|:|:|:|:|  
Db 75 KDKDRDRD 84

RESULT 12  
O16015 PRELIMINARY; PRT; 1034 AA.  
AC O16015;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE EYE COLOR PROTEIN.  
GN GARNET.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
OC Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CANTON S;  
RX MEDLINE; 9744755.  
RA OOI C.E., MOREIRA J.E., DELL'ANGELICA E.C., POY G., WASSARMAN D.A.,  
RA BONIFACINO J.S.;  
RT "Altered expression of a novel adaptin leads to defective pigment  
RT granule biogenesis in the Drosophila eye color mutant garnet.";  
RL EMBO J. 16:4508-4518(1997).  
DR EMBL; AF002164; AAC14585.1; -.  
DR FLYBASE; FBgn001087; g.  
SQ SEQUENCE 1034 AA; 114831 MW; D3FDE8E3 CRC32;

Query Match 81.8%; Score 45; DB 5; Length 1034;  
Best Local Similarity 80.0%; Pred. No. 13;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
      :|:|:|:|:|  
Db 878 KDRKKDKDKD 887

RESULT 13  
Q09655 PRELIMINARY; PRT; 1050 AA.  
ID Q09655;  
AC Q09655;  
DT 01-JAN-1999 (TrEMBLrel. 09, Created)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
DE HYPOTHETICAL 119.3 KD PROTEIN ZK1320.7 IN CHROMOSOME II.  
GN ZK1320.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA BERKS M.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 2 ANK REPEATS.  
EXBL; Z46934; CAA87045.1; -.  
DR WORMPEP; ZK1320.7; CE01705.  
DR PFAM; PF00023; ank; 1.  
KW Hypothetical protein; Repeat; ANK repeat.  
FT DOMAIN 758 761 POLY-ASP.  
FT DOMAIN 826 894 2 X ANK MOTIF REPEATS.  
FT REPEAT 826 858 ANK MOTIF 1.  
FT REPEAT 859 894 ANK MOTIF 2.  
SQ SEQUENCE 1050 AA; 119311 MW; 12409C94 CRC32;

Query Match 81.8%; Score 45; DB 5; Length 1050;  
Best Local Similarity 88.9%; Pred. No. 13;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 9  
      :|:|:|:|:|  
Db 638 KDKDKDKDKD 646

RESULT 14  
O96205 PRELIMINARY; PRT; 2206 AA.  
AC O96205;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 264.1 KD PROTEIN.  
GN PF80560W.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
RT falciparum.";  
RL Science 282:1126-1132(1998).  
DR EMBL; AE001403; AAC71901.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 2206 AA; 264100 MW; AEE134B0 CRC32;

Query Match 81.8%; Score 45; DB 5; Length 2206;  
Best Local Similarity 80.0%; Pred. No. 26;



Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|:|||||:  
Db 666 KDKDKDKNKN 675

RESULT 15

O80716  
ID O80716 PRELIMINARY; PRT: 737 AA.  
AC O80716;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE F14M4.8 PROTEIN.  
GN F14M4.8.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
RT "Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC004411; AAC34219.1; -;  
SQ SEQUENCE 737 AA; 81939 MW; 46858CC9 CRC32;

Query Match 81.8%; Score 45; DB 10; Length 737;

Best Local Similarity 80.0%; Pred. No. 9.7;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KDKDKDKDKD 10  
|:|||||:  
Db 634 KSDKDKDKD 643

Search completed: September 7, 1999, 20:34:32  
Job time: 19728 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:05 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-5  
Perfect score: 55  
Sequence: 1 KFKFKFKFKF 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	1 R71776	Peptide neutralisi
2	55	100.0	10	1 W21593	Antibiotic potenti
3	44	80.0	10	1 W21631	Antibiotic potenti
4	41	74.5	10	1 R71786	Peptide neutralisi
5	41	74.5	10	1 W21603	Antibiotic potenti
6	34	61.8	13	1 R37406	Peptide for treati
7	34	61.8	10	1 R37394	Peptide for treati
8	34	61.8	548	1 W56014	Recombinant botuli
9	34	61.8	1169	1 W56017	Recombinant botuli
10	34	61.8	858	1 W56018	Recombinant botuli
11	34	61.8	177	1 W71504	Helicobacter polyp
12	34	61.8	1291	1 W68392	Clostridium botuli
13	32	58.2	93	1 W20125	H. pylori cell env
14	32	58.2	178	1 W21011	H. pylori cell env
15	32	58.2	388	1 W98618	H. pylori GHP0 231
16	31	56.4	504	1 R24482	Cholesterol 7-alpha
17	31	56.4	10	1 R71783	Peptide neutralisi
18	31	56.4	482	1 R72601	Truncated human ch
19	31	56.4	504	1 R72507	Human cholesterol
20	31	56.4	504	1 R73945	Human cholesterol
21	31	56.4	10	1 W21600	Antibiotic potenti
22	30	54.5	121	1 R29164	PRP3. Isolated DNA
23	30	54.5	151	1 R29165	PRP3 (from genomic
24	30	54.5	649	1 R47504	Pyrococcus furiosu
25	30	54.5	213	1 R47200	Modified xylanase
26	30	54.5	1153	1 R5764	Sequence encoded b
27	30	54.5	28	1 R62493	BPI derived peptid
28	30	54.5	1153	1 R63206	Nitric-oxide-synth
29	30	54.5	28	1 R76336	Bacterial permeabi
30	30	54.5	28	1 R86549	BPI.101 for use in
31	30	54.5	28	1 R81180	Anti-fungal BPI pe
32	30	54.5	28	1 R82381	Bactericidal/perme
33	30	54.5	28	1 R82311	BPI.101, domain II
34	30	54.5	1153	1 R88464	Hepatocyte inducib
35	30	54.5	440	1 R93118	Human GABA-A recep
36	30	54.5	28	1 R78009	BPI (148-161) x 2[
37	30	54.5	28	1 W04093	Antifungal peptide
38	30	54.5	28	1 W05946	Recombinant BPI pe
39	30	54.5	9	1 W07416	Antibacterial/anti
40	30	54.5	543	1 W20785	H. pylori cytoplas
41	30	54.5	440	1 W26464	Human GABA-A recep
42	30	54.5	504	1 W36140	Bovine p58 protein
43	30	54.5	28	1 W43645	Bactericidal/perme

ALIGNMENTS

RESULT 1

R71776  
ID R71776 standard; peptide; 10 AA.  
AC R71776:  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 7; Page 20; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 10 AA:

Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.9e-05;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KFKFKFKFKF 10  
Db 1 KFKFKFKFKF 10  
|||||

RESULT 2

W21593  
ID W21593 standard; peptide; 10 AA.  
AC W21593:  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #5.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 9; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to

44 30 54.5 28 1 W44432 Anti-fungal peptid  
45 30 54.5 1146 1 W96322 Inducible nitric o

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA; 0;

Query Match 100.0%; Score 55; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-05;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
 Db 1 KFKFKFKFK 10  
 |||||

RESULT 3  
 W21631  
 ID W21631 standard; peptide; 10 AA.  
 AC W21631;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #43.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..10  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 47; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA; 0;

Query Match 80.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.006;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 8  
 Db 2 KFKFKFKFK 9  
 |||||

RESULT 4  
 R71786  
 ID R71786 standard; peptide; 10 AA.  
 AC R71786;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.

PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 17; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 74.5%; Score 41; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.021;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 9  
 Db 1 KFKFKFKFK 9  
 |||||

RESULT 5  
 W21603  
 ID W21603 standard; peptide; 10 AA.  
 AC W21603;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #15.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 19; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 74.5%; Score 41; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.021;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 9  
 Db 1 KFKFKFKFK 9  
 |||||

RESULT 6  
 ID R57406 standard; Protein; 13 AA.  
 AC R57406;  
 DT 21-MAR-1995 (first entry)  
 DE Peptide for treating diseases related to anti-DNA antibodies.  
 KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
 OS Synthetic.  
 PN J06192290-A.  
 PD 12-JUL-1994.  
 PF 18-JAN-1993; 006098.  
 PR 30-SEP-1992; JP-261821.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 94-260510/32.  
 PT A peptide and an adsorbing agent prep'd. by immobilising it on a  
 PT carrier - useful for treatment of diseases related to anti-DNA  
 PT antibodies and immune complexes  
 PS Disclosure; Page 13; 14pp; Japanese.  
 CC The sequences given in R57386-413 are peptides which are all covered  
 CC by the claimed generic formula:  
 CC H-X-(A-B)n-Y-Z  
 CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
 CC B = Trp, Phe, Asn or Glu;  
 CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
 CC consisting of 2-10 residues, provided that at least one of  
 CC X or Y are present;  
 CC Z = OH or NH2; and  
 CC n = 2-5.  
 CC These peptides may be immobilised on a carrier in the preparation of an  
 CC adsorbing agent which may be used in the treatment of diseases related  
 CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 13 AA;

Query Match 61.8%; Score 34; DB 1; Length 13;  
 Best Local Similarity 60.0%; Pred. No. 0.52;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKFKEFKFKF 10  
 | | | | | | | |  
 Db 3 KKKEFEFEF 12

RESULT 7  
 ID R57394 standard; Protein; 10 AA.  
 AC R57394;  
 DT 21-MAR-1995 (first entry)  
 DE Peptide for treating diseases related to anti-DNA antibodies.  
 KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
 OS Synthetic.  
 PN J06192290-A.  
 PD 12-JUL-1994.  
 PF 18-JAN-1993; 006098.  
 PR 30-SEP-1992; JP-261821.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 94-260510/32.  
 PT A peptide and an adsorbing agent prep'd. by immobilising it on a  
 PT carrier - useful for treatment of diseases related to anti-DNA  
 PT antibodies and immune complexes  
 PS Disclosure; Page 12; 14pp; Japanese.  
 CC The sequences given in R57386-413 are peptides which are all covered  
 CC by the claimed generic formula:  
 CC H-X-(A-B)n-Y-Z  
 CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
 CC B = Trp, Phe, Asn or Glu;  
 CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
 CC consisting of 2-10 residues, provided that at least one of  
 CC X or Y are present;  
 CC Z = OH or NH2; and  
 CC n = 2-5.  
 CC These peptides may be immobilised on a carrier in the preparation of an  
 CC adsorbing agent which may be used in the treatment of diseases related

CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 10 AA;

Query Match 61.8%; Score 34; DB 1; Length 10;  
 Best Local Similarity 55.8%; Pred. No. 0.4;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | | | | | | | |  
 Db 1 FEFEFEFEF 9

RESULT 8  
 ID W56014 standard; Protein; 548 AA.  
 AC W56014;  
 DT 27-JUL-1998 (first entry)  
 DE Recombinant botulinum neurotoxin type B LH107/B.  
 KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;  
 KW Immunogen; detection; tetanus; non-toxic; toxin.  
 OS Synthetic.  
 OS Clostridium botulinum.  
 PN W09807864-A1.  
 PD 26-FEB-1998.  
 PF 22-AUG-1997; G02273.  
 PR 13-DEC-1996; GB-025996.  
 PR 23-AUG-1996; GB-017671.  
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
 PA (SPEX-) SPEYWOOD LAB LTD.  
 PI Foster KA, Quinn CP, Shone CC;  
 DR WPI: 98-169168/15.  
 DR N-PSDB: V26286.  
 PT Recombinant polypeptides - used to develop therapeutic  
 PT agents, immunogens or as non-toxic standards for the detection of  
 PT neurotoxins  
 PS Example 2; Page 103-104; 137pp; English.  
 CC The present sequence represents a recombinant neurotoxin protein from  
 CC the present invention. The present invention describes recombinant  
 CC neurotoxin proteins which comprise a first and second domain, where  
 CC the first domain is adapted to cleave one or more vesicle or  
 CC plasma-membrane associated proteins essential to exocytosis, and where  
 CC the second domain is adapted: (a) to translocate the protein into a  
 CC cell; (b) to increase the solubility of the protein compared to the  
 CC solubility of the first domain on its own, or (c) both to translocate  
 CC the protein into a cell and to increase the solubility of the protein  
 CC compared to the solubility of the first domain on its own, the protein  
 CC being free of clostridial neurotoxin (CN) and free of CN precursor that  
 CC can be converted into toxin by proteolytic action. The recombinant  
 CC proteins can be used as therapeutic agents for targeting cells  
 CC expressing a relevant substrate. The products can also be used as  
 CC immunogens and as non-toxic standards for the assessment and development  
 CC of in vitro assays for the detection of functional botulinum or tetanus  
 CC neurotoxins either in foodstuffs or in environmental samples.  
 SQ Sequence 548 AA;

Query Match 61.8%; Score 34; DB 1; Length 548;  
 Best Local Similarity 70.0%; Pred. No. 21;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKFKEFKFKF 10  
 | | | | | | | |  
 Db 321 KKFKEFKFKF 330

RESULT 9  
 ID W56017 standard; Protein; 1169 AA.  
 AC W56017;  
 DT 27-JUL-1998 (first entry)  
 DE Recombinant botulinum neurotoxin type B LH728/B.  
 KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;

KW immunogen; detection; tetanus; non-toxic; toxin.  
 OS Synthetic.  
 OS Clostridium botulinum.  
 PN W09807864-A1.  
 PD 26-FEB-1998.  
 PF 22-AUG-1997; G02273.  
 PR 13-DEC-1996; GB-025996.  
 PR 23-AUG-1996; GB-017671.  
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
 PA (SPEY-) SPEYWOOD LAB LTD.  
 PI Foster KA, Quinn CP, Shone CC;  
 DR WPI: 98-169168/15.  
 DR N-PSDB: V26289.  
 PT Recombinant neurotoxin polypeptides - used to develop therapeutic  
 PT agents, immunogens or as non-toxic standards for the detection of  
 PT neurotoxins  
 FS Example 2: Page 91-94; 137pp; English.  
 CC The present sequence represents a recombinant neurotoxin protein from  
 CC the present invention. The present invention describes recombinant  
 CC neurotoxin proteins which comprise a first and second domain, where  
 CC the first domain is adapted to cleave one or more vesicle or  
 CC plasma-membrane associated proteins essential to exocytosis, and where  
 CC the second domain is adapted: (a) to translocate the protein into a  
 CC cell; (b) to increase the solubility of the protein compared to the  
 CC solubility of the first domain on its own, or (c) both to translocate  
 CC the protein into a cell and to increase the solubility of the protein  
 CC compared to the solubility of the first domain on its own, the protein  
 CC being free of clostridial neurotoxin (CN) and free of CN precursor that  
 CC can be converted into toxin by proteolytic action. The recombinant  
 CC proteins can be used as therapeutic agents for targeting cells  
 CC expressing a relevant substrate. The products can also be used as  
 CC immunogens and as non-toxic standards for the assessment and development  
 CC of in vitro assays for the detection of functional botulinum or tetanus  
 CC neurotoxins either in foodstuffs or in environmental samples.  
 SQ Sequence 1169 AA;

Query Match 61.8%; Score 34; DB 1; Length 1169;  
 Best Local Similarity 70.0%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKFKF 10  
 | | | | |  
 DB 321 KNKFKDKYKF 330

## RESULT 10

ID W56018 standard; Protein; 858 AA.  
 AC W56018;  
 DT 27-JUL-1998 (first entry)  
 DE Recombinant botulinum neurotoxin type B LH417/B.  
 KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;  
 KW immunogen; detection; tetanus; non-toxic; toxin.  
 OS Synthetic.  
 OS Clostridium botulinum.  
 PN W09807864-A1.  
 PD 26-FEB-1998.  
 PF 22-AUG-1997; G02273.  
 PR 13-DEC-1996; GB-025996.  
 PR 23-AUG-1996; GB-017671.  
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
 PA (SPEY-) SPEYWOOD LAB LTD.  
 PI Foster KA, Quinn CP, Shone CC;  
 DR WPI: 98-169168/15.  
 DR N-PSDB: V26290.  
 PT Recombinant neurotoxin polypeptides - used to develop therapeutic  
 PT agents, immunogens or as non-toxic standards for the detection of  
 PT neurotoxins  
 FS Example 2: Page 98-100; 137pp; English.  
 CC The present sequence represents a recombinant neurotoxin protein from  
 CC the present invention. The present invention describes recombinant  
 CC neurotoxin proteins which comprise a first and second domain, where

CC the first domain is adapted to cleave one or more vesicle or  
 CC plasma-membrane associated proteins essential to exocytosis, and where  
 CC the second domain is adapted: (a) to translocate the protein into a  
 CC cell; (b) to increase the solubility of the protein compared to the  
 CC solubility of the first domain on its own, or (c) both to translocate  
 CC the protein into a cell and to increase the solubility of the protein  
 CC compared to the solubility of the first domain on its own, the protein  
 CC being free of clostridial neurotoxin (CN) and free of CN precursor that  
 CC can be converted into toxin by proteolytic action. The recombinant  
 CC proteins can be used as therapeutic agents for targeting cells  
 CC expressing a relevant substrate. The products can also be used as  
 CC immunogens and as non-toxic standards for the assessment and development  
 CC of in vitro assays for the detection of functional botulinum or tetanus  
 CC neurotoxins either in foodstuffs or in environmental samples.  
 SQ Sequence 858 AA;

Query Match 61.8%; Score 34; DB 1; Length 858;  
 Best Local Similarity 70.0%; Pred. No. 33;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKFKF 10  
 | | | | |  
 DB 321 KNKFKDKYKF 330

## RESULT 11

W71504  
 ID W71504 standard; Protein; 177 AA.  
 AC W71504;  
 DT 09-NOV-1998 (first entry)  
 DE Helicobacter polypeptide GHPO 1200.  
 KW GHPO 1200; infection; therapy; diagnosis; vaccine; gastritis;  
 KW ulcer.  
 OS Helicobacter pylori.  
 PN W09821225-A1.  
 PD 22-MAY-1998.  
 PF 14-NOV-1997; U21353.  
 PR 29-JUL-1997; US-902615.  
 PR 14-NOV-1996; US-749051.  
 PR 01-APR-1997; US-831309.  
 PR 01-APR-1997; US-833457.  
 PR 01-APR-1997; US-834705.  
 PR 24-JUN-1997; US-881227.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 PI Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C,  
 PI Odenbreit S, Tomb J;  
 DR WPI: 98-297855/26.  
 DR N-PSDB: V52039.  
 PT Helicobacter polynucleotide and polypeptide sequences - useful to  
 PT treat or prevent gastrointestinal infection  
 PS Claim 1; Page 162-163; 362pp; English.  
 CC This claimed Helicobacter pylori polypeptide, designated GHPO 1200,  
 CC can be used in vaccination methods for preventing or treating  
 CC Helicobacter infection. 85 Helicobacter polypeptides (see  
 CC W71474-W71558) are claimed, as well as isolated polynucleotides  
 CC (see V52009-93) that encode them. The invention also provides:  
 CC methods for producing these Helicobacter polypeptides in  
 CC recombinant host systems, and related expression cassettes, vectors  
 CC and transformed or transfected host cells; live vaccine vectors  
 CC that contain the polynucleotides of the invention and which can be  
 CC used to prevent or treat Helicobacter infection; therapeutic and/or  
 CC prophylactic methods involving administration of polynucleotide  
 CC molecules, polypeptides or monospecific antibodies; methods for  
 CC detecting the presence of Helicobacter in samples using e.g.  
 CC the polypeptides or monospecific antibodies; and methods for  
 CC purifying the polypeptides by antibody-based affinity  
 CC chromatography.  
 SQ Sequence 177 AA;

Query Match 61.8%; Score 34; DB 1; Length 177;  
 Best Local Similarity 60.0%; Pred. No. 6.9;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KFKFKFKFKF 10  
 : | | | | |  
 DB 88 RFSFPFVKF 97

## RESULT 12

W68392 ID W68392 standard; Protein; 1291 AA.  
 AC W68392; 1998 (first entry)  
 DT Clostridium botulinum type B toxin.  
 DE Antitoxin; vaccine; neurotoxin; toxin B; intoxication; immunogen;  
 KW botulism; BotB.  
 OS Clostridium botulinum serotype B Danish strain.  
 PN W09808540-A1.  
 PD 05-MAR-1998.  
 PF 28-AUG-1997; U15394.  
 PR 28-AUG-1996; US-704159.  
 PA (OPHI-) OPHIDIAN PHARM INC.  
 PI Thalley BS, Williams JA;  
 DR WPI; 98-230234/20.  
 DR N-PSDB; V30579.  
 PT Host cell containing recombinant expression vector encoding  
 PT Clostridium botulinum type B or E toxin - useful to treat humans  
 PT and other animals at risk of intoxication with clostridial toxin  
 PS Example 35; Page 291-296; 428pp; English.  
 CC This is the amino acid sequence of the type B toxin of Clostridium  
 CC botulinum Danish strain. The C fragment (see W68393-94) of the  
 CC serotype B toxin has been expressed as a histidine-tagged protein  
 CC in Escherichia coli. The invention relates to C. botulinum  
 CC recombinant toxins. Methods are provided which allow for the  
 CC isolation of soluble recombinant proteins free of significant  
 CC endotoxin contamination. Preferred hosts for production of the  
 CC recombinant proteins are E. coli, insect cells and yeast cells.  
 CC The recombinant proteins are used as immunogens for the production  
 CC of vaccines and antitoxins that are useful in the treatment of  
 CC humans and animals at risk of intoxication with clostridial toxin.  
 CC Sequence 1291 AA;

Query Match 61.8%; Score 34; DB 1; Length 1291;  
 Best Local Similarity 70.0%; Pred. No. 49;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KFKFKFKFKF 10  
 : | | | | |  
 DB 321 KNKFKDKYKF 330

## RESULT 13

W20125 ID W20125 standard; Protein; 93 AA.  
 AC W20125;  
 DT 07-JUL-1997 (first entry)  
 DE H. pylori cell envelope protein 1365943.aa.  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 OS Helicobacter pylori.  
 PN W09640893-A1.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09122.  
 PR 07-JUN-1995; US-487032.  
 PR 01-APR-1996; US-630405.  
 PA (ASTR ) ASTRA AB.  
 PI Berglindh OT, Smith D, Melligaard BL;  
 DR WPI; 97-052306/05.  
 DR N-PSDB; T67370.  
 PT Helicobacter pylori nucleic acid sequences and related

PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 PS Claim 56; Page 353; 1481pp; English.  
 CC The present sequence shows a Helicobacter pylori cell envelope  
 CC protein which may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 93 AA;

Query Match 58.2%; Score 32; DB 1; Length 93;  
 Best Local Similarity 60.0%; Pred. No. 8.4;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KFKFKFKFKF 10  
 : | | | | |  
 DB 82 RFGFPFVKF 91

## RESULT 14

W21011 ID W21011 standard; protein; 178 AA.  
 AC W21011;  
 DT 21-JUL-1997 (first entry)  
 DE H. pylori cell envelope protein, hp5el2982orf14.  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 OS Helicobacter pylori.  
 PN W09640893-A1.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09122.  
 PR 07-JUN-1995; US-487032.  
 PR 01-APR-1996; US-630405.  
 PA (ASTR ) ASTRA AB.  
 PI Berglindh OT, Smith D, Melligaard BL;  
 DR WPI; 97-052306/05.  
 DR N-PSDB; T68264.  
 PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 PS Claim 56; Pages 1400; 1481pp; English.

CC The present sequence is a Helicobacter pylori cell envelope protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds, useful  
 CC as potential H. pylori life cycle activators or inhibitors. The genomic  
 CC sequence of H. pylori (ATCC 55679) was determined from overlapping  
 CC contigs generated by mechanically shearing the bacterial DNA. The  
 CC sequences were analysed for ORF of at least 180 nucleotides, and the  
 CC predicted coding regions defined by computer evaluation. To identify  
 CC likely H. pylori antigens for vaccine development, the amino acid  
 CC sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 178 AA;

Query Match 58.2%; Score 32; DB 1; Length 178;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKFKKKF 10  
: | | | |  
Db 89 RGGFFFKF 98

RESULT 15  
W98618  
ID W98618 standard; Protein; 388 AA.  
AC W98618;  
DT 31-MAR-1999 (first entry)  
DE H. pylori GHPO 231 protein.  
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
KW peptic ulcer disease.  
OS Helicobacter pylori.  
PN W09843478-Al.  
PD 08-OCT-1998.  
PF 01-APR-1998; U06371.  
PR 29-JUL-1997; US-902615.  
PR 01-APR-1997; US-833457.  
PR 24-JUN-1997; US-881227.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
DR WPI: 98-542293/46.  
DR N-PSDB: X1437.  
PT New isolated Helicobacter polynucleotides - used to develop products  
PT for the diagnosis, prevention and treatment of Helicobacter  
PT infections and gastrointestinal diseases  
PS Claim 8; Page 1249-1251; 2054pp; English.  
CC This sequence represents a Helicobacter pylori GHPO protein of the  
CC invention. The polypeptides can be used for preventing or treating  
CC Helicobacter infections, and gastroduodenal diseases associated with  
CC these infections, including acute, chronic, and atrophic gastritis, and  
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
CC used for the production of antibodies. The products can also be used for  
CC detection and diagnosis.  
SQ Sequence 388 AA;

Query Match 58.2%; Score 32; DB 1; Length 388;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKFKKKF 10  
: | | | |  
Db 166 KKKTFHFSF 175

Search completed: September 7, 1999, 20:37:07  
Job time: 18456 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:32 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-5  
Perfect score: 55  
Sequence: 1 KFKFKFKFKF 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	1	US-08-097-830E-5
2	55	100.0	10	2	US-08-456-112B-5
3	44	80.0	10	2	US-08-456-112B-43
4	41	74.5	10	1	US-08-097-830E-15
5	41	74.5	10	2	US-08-456-112B-15
6	31	56.4	482	1	US-08-135-510-5
7	31	56.4	504	1	US-08-135-511-28
8	31	56.4	482	1	US-08-483-852-5
9	31	56.4	504	1	US-08-483-852-8
10	31	56.4	10	1	US-08-097-830E-12
11	31	56.4	504	1	US-08-361-458-3
12	31	56.4	482	1	US-08-477-953-5
13	31	56.4	504	1	US-08-477-953-8
14	31	56.4	504	2	US-08-187-453-28
15	31	56.4	10	2	US-08-456-112B-12
16	31	56.4	482	2	US-08-477-952-5
17	31	56.4	504	2	US-08-477-952-8
18	30	54.5	213	1	US-08-104-445-3
19	30	54.5	649	1	US-07-894-212A-2
20	30	54.5	647	1	US-07-894-212A-8
21	30	54.5	1153	1	US-08-314-917-2
22	30	54.5	28	1	US-08-311-611A-95
23	30	54.5	650	1	US-07-893-928A-1
24	30	54.5	28	1	US-08-372-783-95
25	30	54.5	28	1	US-08-372-105-95
26	30	54.5	28	1	US-08-306-473A-95
27	30	54.5	440	1	US-08-459-100A-2
28	30	54.5	1153	1	US-08-265-046-2
29	30	54.5	28	1	US-08-209-762-95
30	30	54.5	28	2	US-08-473-344-95
31	30	54.5	28	2	US-08-621-803-89
32	30	54.5	28	2	US-08-485-445A-95
33	30	54.5	28	2	US-08-621-259A-33
34	30	54.5	1153	2	US-08-465-522-2
35	30	54.5	1153	3	PCT-US93-11401-2
36	30	54.5	28	3	PCT-US94-02465-95
37	30	54.5	440	3	PCT-US94-09589-2
38	30	54.5	28	3	PCT-US95-00498-95
39	30	54.5	28	3	PCT-US95-00656-95

40 30 54.5 1153 3 PCT-US95-07849-2 Sequence 2, Appl  
41 30 54.5 28 3 PCT-US95-09262-33 Sequence 33, Appl  
42 29 52.7 170 1 US-08-127-499A-20 Sequence 20, Appl  
43 29 52.7 102 1 US-08-169-701-1 Sequence 1, Appl  
44 29 52.7 170 1 US-08-482-847-20 Sequence 20, Appl  
45 29 52.7 268 2 US-08-207-412B-12 Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-5  
; Sequence 5, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-097-830E-5  
  
Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KFKFKFKFKF 10  
Db 1 KFKFKFKFKF 10  
  
RESULT 2  
US-08-456-112B-5  
; Sequence 5, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan

Wed Sep 8 09:45:34 1999

```
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: LEADING EDGE 486
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456.112B
FILING DATE: May 31, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-456-112B-5

Query Match      100.0%; Score 55; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10
Db 1 KFKFKFKFK 10

RESULT 3
US-08-456-112B-43
; Sequence 43, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Poirio, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; OTHER INFORMATION: sulfide bond between Cys and
; OTHER INFORMATION: Cys
US-08-097-830E-15
; Sequence 15, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Poirio, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; OTHER INFORMATION: sulfide bond between Cys and
; OTHER INFORMATION: Cys
US-08-097-830E-15

Query Match      74.5%; Score 41; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.042;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 9
Db 1 KFKCKFKFK 9
```

```
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: LEADING EDGE 486
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456.112B
FILING DATE: May 31, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Costigan, James V.
REGISTRATION NUMBER: 25,669
REFERENCE/DOCKET NUMBER: 576-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 302-8989
TELEFAX: (212) 302-8998
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: circular
US-08-456-112B-5

Query Match      100.0%; Score 55; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10
Db 1 KFKFKFKFK 10

RESULT 3
US-08-456-112B-43
; Sequence 43, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Poirio, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
```

RESULT 5  
US-08-456-112B-15  
; Sequence 15, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-15

Query Match 74.5%; Score 41; DB 2; Length 10;  
Best Local Similarity 88.9%; Pred. NO. 0.042;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKFKFK 9  
Db 1 KFKFKFK 9

RESULT 6  
US-08-135-510-5  
; Sequence 5, Application US/08135510  
; Patent No. 5420028  
; GENERAL INFORMATION:  
; APPLICANT: CHIANG, John Young Ling  
; TITLE OF INVENTION: Truncated Human Cholesterol  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/135.510  
; FILING DATE: 13-OCT-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/176 HOCE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 482 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-135-510-5

Query Match 56.4%; Score 31; DB 1; Length 482;  
Best Local Similarity 71.4%; Pred. No. 63;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFKFKFK 7  
Db 474 EFKYFK 480

RESULT 7  
US-08-135-511-28  
; Sequence 28, Application US/08135511  
; Patent No. 5558999  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, John  
; TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene  
; TITLE OF INVENTION: Regulatory Elements and Methods for Using Them  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/135.511  
; FILING DATE: 13-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/175  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-135-511-28

Query Match 56.4%; Score 31; DB 1; Length 504;  
Best Local Similarity 71.4%; Pred. No. 66;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFK 7  
:|:|:|  
Db 496 EFKYKFK 502

RESULT 8  
US-08-483-852-5  
; Sequence 5, Application US/08483852  
; Patent No. 5650286  
; GENERAL INFORMATION:  
; APPLICANT: CHIANG, John Young Ling  
; TITLE OF INVENTION: Genomic DNA of Human Cholesterol  
; TITLE OF INVENTION: 7a-Hydroxylase and Methods of Using It  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/483,852  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; PRIOR APPLICATION NUMBER: US 08/361,458  
; FILING DATE: 21-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,511  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,488  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,510  
; FILING DATE: 13-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/220 HOCE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 482 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-483-852-5

Query Match 56.4%; Score 31; DB 1; Length 482;  
Best Local Similarity 71.4%; Pred. No. 63;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFK 7  
:|:|:|  
Db 474 EFKYKFK 480

RESULT 9  
US-08-483-852-8  
; Sequence 8, Application US/08483852  
; Patent No. 5650286  
; GENERAL INFORMATION:  
; APPLICANT: CHIANG, John Young Ling

; TITLE OF INVENTION: Genomic DNA of Human Cholesterol  
; TITLE OF INVENTION: 7a-Hydroxylase and Methods of Using It  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,852  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/361,458  
; FILING DATE: 21-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,511  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,488  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,510  
; FILING DATE: 13-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/220 HOCE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-483-852-8

Query Match 56.4%; Score 31; DB 1; Length 504;  
Best Local Similarity 71.4%; Pred. No. 66;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFK 7  
:|:|:|  
Db 496 EFKYKFK 502

RESULT 10  
US-08-097-830E-12  
; Sequence 12, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-12

Query Match 56.4%; Score 31; DB 1; Length 10;  
Best Local Similarity 77.8%; Pred. No. 2;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKKKKKKK 9  
Db 2 KKKKKKKK 10

RESULT 11  
US-08-361-458-3  
Sequence 3, Application US/08361458  
Patent No. 5663483  
GENERAL INFORMATION:  
APPLICANT: Chiang, John  
TITLE OF INVENTION: Genomic DNA of Human Cholesterol  
OPERATING SYSTEM: PC-DOS/MS-DOS  
FILING DATE: 21-DEC-1994  
NUMBER OF SEQUENCES: 6  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,458  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/135,488  
FILING DATE: 13-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SANDERCOCK, Colin G.  
REGISTRATION NUMBER: 31,298  
REFERENCE/DOCKET NUMBER: 18748/174  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 504 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-361-458-3

Query Match 56.4%; Score 31; DB 1; Length 504;  
Best Local Similarity 71.4%; Pred. No. 66;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKK 7  
Db 496 EFKYKFK 502

RESULT 12  
US-08-477-953-5  
Sequence 5, Application US/08477953  
Patent No. 5677159  
GENERAL INFORMATION:  
APPLICANT: CHIANG, John Young Ling  
TITLE OF INVENTION: Genomic DNA of Human Cholesterol  
OPERATING SYSTEM: PC-DOS/MS-DOS  
FILING DATE: 21-DEC-1994  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,953  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/361,458  
FILING DATE: 21-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/135,511  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/135,488  
FILING DATE: 13-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/135,510  
FILING DATE: 13-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SANDERCOCK, Colin G.  
REGISTRATION NUMBER: 31,298  
REFERENCE/DOCKET NUMBER: 18748/219 HOCE  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 482 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-477-953-5

Query Match 56.4%; Score 31; DB 1; Length 482;  
Best Local Similarity 71.4%; Pred. No. 63;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFKFKFK 7  
:||:||||  
Db 474 EFKYKFK 480

## RESULT 13

US-08-477-953-8  
; Sequence 8, Application US/08477953  
; Patent No. 5677159  
; GENERAL INFORMATION:  
; APPLICANT: CHIANG, John Young Ling  
; TITLE OF INVENTION: Genomic DNA of Human Cholesterol  
; TITLE OF INVENTION: 7a-Hydroxylase and Methods of Using It  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,953  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/361,458  
; FILING DATE: 21-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,511  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,488  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,510  
; FILING DATE: 13-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/219 HOCE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-477-953-8

Query Match 56.4%; Score 31; DB 1; Length 504;  
Best Local Similarity 71.4%; Pred. No. 66;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFKFKFK 7  
:||:||||  
Db 496 EFKYKFK 502

## RESULT 14

US-08-187-453-28  
; Sequence 28, Application US/08187453  
; Patent No. 5753431  
; GENERAL INFORMATION:

; APPLICANT: Chiang, John  
; TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene  
; TITLE OF INVENTION: Regulatory Elements and Transcription Factors  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/187,453  
; FILING DATE: 28-JAN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,488  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,511  
; FILING DATE: 13-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/135,510  
; FILING DATE: 13-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 18748/188  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-187-453-28

Query Match 56.4%; Score 31; DB 2; Length 504;  
Best Local Similarity 71.4%; Pred. No. 66;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFKFKFK 7  
:||:||||  
Db 496 EFKYKFK 502

## RESULT 15

US-08-456-112B-12  
; Sequence 12, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486

; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-456-112B-12

Query Match 56.4%; Score 31; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 9
Db 2 KKKKKFKFK 10

Search completed: September 7, 1999, 22:38:32
Job time: 7916 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:17 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-5

Perfect score: 55

Sequence: 1 KFKFKFKFKF 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	63.6	434	2	S61999
2	35	63.6	434	2	S55155
3	34	61.8	488	1	QOBHS
4	34	61.8	1291	1	A48940
5	34	61.8	1291	2	I40631
6	34	61.8	177	2	D64523
7	34	61.8	348	2	A71153
8	33	60.0	266	2	A25157
9	33	60.0	266	2	S55397
10	33	60.0	2470	2	S57085
11	33	60.0	378	2	I40610
12	33	60.0	609	2	S36569
13	33	60.0	669	2	S62819
14	33	60.0	626	2	A70169
15	33	60.0	1142	2	A45031
16	33	60.0	1175	2	S52417
17	32	58.2	631	2	G70188
18	32	58.2	209	2	H70189
19	32	58.2	881	2	S46633
20	32	58.2	175	2	F71982
21	32	58.2	516	2	S34525
22	32	58.2	660	2	F70101
23	32	58.2	660	2	G64614
24	32	58.2	417	2	H71898
25	32	58.2	469	2	S74825
26	32	58.2	191	2	T01132
27	32	58.2	1181	2	T01799
28	32	58.2	337	2	S56017
29	32	58.2	337	2	S53846
30	31	56.4	504	1	JH0659
31	31	56.4	109	1	MNIH82
32	31	56.4	501	2	I46701
33	31	56.4	137	2	S68429
34	31	56.4	1146	2	I64112
35	31	56.4	1036	2	H64245
36	31	56.4	728	2	G54245
37	31	56.4	628	2	A69381
38	31	56.4	1146	2	E70204
39	31	56.4	673	2	S73444

40 31 56.4 190 2 C69345 hypothetical prote  
41 31 56.4 2136 2 A05037 hypothetical prote  
42 31 56.4 601 2 JE0238 stress protein p66  
43 31 56.4 261 2 S51994 hypothetical prote  
44 31 56.4 294 2 S60991 hypothetical prote  
45 31 56.4 1029 2 S64731 serine/threonine-s

ALIGNMENTS

RESULT 1

S61999

hypothetical protein YPL123c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein LPH4c

C:Species: Saccharomyces cerevisiae

C>Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 12-Dec-1997

C:Accession: S61999

R:Schlenstedt, G.; Silver, P.A.

submitted to the EMBL Data Library, December 1995

A:Reference number: S61996

A:Accession: S61999

A:Molecule type: DNA

A:Residues: 1-434 <SCH>

A:Cross-references: EMBL:U43503; NID:gl163087; PID:gl163091; MIPS:YPL123c

C:Genetics:

A:Map position: 16L

Query Match 63.6% Score 35; DB 2; Length 434;  
Best Local Similarity 75.0%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKFKFKF 8

Db 424 KYKFKLKF 431

RESULT 2

S55155

probable tyrosine phosphatase YNL128w - yeast (Saccharomyces cerevisiae)

N:Alternate names: probable tyrosine phosphatase N1220; protein JTB434; protein N1872

C:Species: Saccharomyces cerevisiae

C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 06-Feb-1998

C:Accession: S55155; S59260; S63073

R:Mallet, L.; Bussereau, F.; Jacquet, M.

submitted to the EMBL Data Library, November 1994

A:Description: A 43.5 kb fragment of the chromosome XIV.

A:Reference number: S55136

A:Accession: S55155

A:Molecule type: DNA

A:Residues: 1-434 <NAL>

A:Cross-references: EMBL:Z46843; NID:g861113; PID:g854509

R:Mallet, L.; Bussereau, F.; Jacquet, M.

Yeast 11, 1195-1209, 1995

A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MFA2, MEP2, CAP/SR

A:Reference number: S59241; MUID:96109932

A:Accession: S59260

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-434 <NAW>

A:Cross-references: EMBL:Z46843; NID:g861113; PID:g854509

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994

R:Mallet, L.; Bussereau, F.; Jacquet, M.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63069

A:Accession: S63073

A:Molecule type: DNA

A:Residues: 1-434 <NAF>

A:Cross-references: EMBL:Z71404; NID:gl302063; PID:e239804; MIPS:YNL128

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:TEP1

A:Cross-references: SGD:S0005072; MIPS:YNL128w  
 A:Map position: 14L  
 C:Keywords: transmembrane protein  
 F:38-54/Domain: transmembrane #status predicted <TML>

Query Match 63.6%; Score 35; DB 2; Length 434;  
 Best Local Similarity 85.7%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKFKFK 7  
 | | | | |  
 DB 268 KFKFKFK 274

## RESULT 3

QOBEHS  
 alkaline exonuclease (EC 3.1.11.-) - human herpesvirus 6 (strain Uganda-1102)  
 C:Species: human herpesvirus 6  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 24-Oct-1997  
 C:Accession: F36769  
 R:Lawrence, G.L.; Chee, M.; Craxton, M.A.; Compels, U.A.; Honess, R.W.; Barrell, B.G.  
 J. Virol. 64, 287-299, 1990  
 A:Title: Human herpesvirus 6 is closely related to human cytomegalovirus.  
 A:Reference number: A33560; MUID:90080132  
 A:Accession: F36769  
 A:Molecule type: DNA  
 A:Residues: 1-488 <LAW>  
 A:Cross-references: GB:M68963; GB:M28243; NID:g325494; PID:g325510  
 C:Superfamily: human cytomegalovirus alkaline exonuclease  
 C:Keywords: exonuclease; hydrolase

Query Match 61.8%; Score 34; DB 1; Length 488;  
 Best Local Similarity 55.6%; Pred. No. 30;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFKFKFKF 10  
 | | | | |  
 DB 235 FEIKFKYK 243

## RESULT 4

A48940  
 botoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum  
 N:Alternate names: botulinum neurotoxin type B (BoNT/B)  
 C:Species: Clostridium botulinum  
 C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-May-1998  
 C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574  
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.  
 Appl. Environ. Microbiol. 58, 2345-2354, 1992  
 A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B precursor  
 A:Reference number: A48940; MUID:92384550  
 A:Accession: A48940  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <WHE>  
 A:Cross-references: GB:M81186; NID:g144734; PID:g144735  
 A:Experimental source: type B, Danish  
 A:Note: Sequence extracted from NCBI backbone (NCBI:112080, NCBI:P:112081); this publication  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J. Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific isoforms  
 A:Reference number: S48103; MUID:94013372  
 A:Accession: S48105  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-994 <CAM>  
 A:Cross-references: EMBL:X70817; NID:g407782; PID:g407783  
 A:Experimental source: proteolytic type B, strain NCTC 7273  
 R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.  
 Submitted to the EMBL Data Library, April 1992  
 A:Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison

A:Reference number: S21575  
 A:Accession: S21575  
 A:Molecule type: DNA  
 A:Residues: 36-217, 'G', '219-224', 'S', '226-246' <SZA>  
 A:Cross-references: EMBL:211934; NID:g40383; PID:g40384  
 R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars  
 J. Biol. Chem. 267, 14721-14729, 1992  
 A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus  
 A:Reference number: A42871; MUID:92340509  
 A:Accession: A42871  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-313, 'S', '315-451' <KUR>  
 A:Experimental source: strain Okra  
 A:Note: sequence extracted from NCBI backbone (NCBI:109365)  
 R:DasGupta, B.R.; Datta, A.  
 Biochimie 70, 811-817, 1988  
 A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with  
 A:Reference number: S07155; MUID:89000987  
 A:Accession: S07155  
 A:Molecule type: protein  
 A:Residues: 2-29, 'M', '31-45' <DAS>  
 A:Accession: S08562  
 A:Molecule type: protein  
 A:Residues: 442-463, 'R', '465-467' <DA2>  
 R:Schmidt, J.J.; Sathymoorthy, V.; DasGupta, B.R.  
 Arch. Biochem. Biophys. 238, 544-548, 1985  
 A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.  
 A:Reference number: S07128; MUID:85197963  
 A:Accession: S07128  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-16 <SCH1>  
 A:Accession: S08573  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 2-17 <SCH2>  
 A:Accession: S08574  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 442-459 <SCH3>  
 R:Schiaivo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.  
 Nature 359, 832-835, 1992  
 A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo  
 A:Reference number: S21125; MUID:93063293  
 A:Contents: annotation  
 C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy  
 C:Genetics:  
 A:Gene: bont/b  
 C:Function:  
 A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2  
 C:Superfamily: tetanus toxin  
 C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
 F:2-441/Product: bontoxilysin B light chain #status experimental <LGH>  
 F:442-1291/Product: bontoxilysin B heavy chain #status experimental <HVI>  
 F:230,234/Binding site: zinc (His) #status predicted  
 F:231/Active site: Glu #status predicted

Query Match 61.8%; Score 34; DB 1; Length 1291;  
 Best Local Similarity 70.0%; Pred. No. 76;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
 | | | | |  
 DB 321 KNKFKDKFK 330

## RESULT 5

I40631  
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum  
 C:Species: Clostridium botulinum  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 05-Jun-1998

C:Accession: I40631; S48103; S48104; S36015  
 R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.  
 Curr. Microbiol. 28, 101-110, 1994  
 A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum  
 A:Reference number: I40631; MUID:941122659  
 A:Accession: I40631  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1291 <RES>  
 A:Cross-references: EMBL:X71343; NID:9296148; PID:g296149  
 R:Campbell, K.D.; Collins, M.D.; East, A.K.  
 J. Clin. Microbiol. 31, 2255-2262, 1993  
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id  
 A:Reference number: S48103; MUID:94013372  
 A:Accession: S48103  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>  
 A:Cross-references: EMBL:X70814; NID:9407778; PID:g407779  
 A:Experimental source: non-proteolytic strain 2129B (Scott)  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
 A:Accession: S48104  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>  
 A:Cross-references: EMBL:X70819; NID:9407780; PID:g407781  
 A:Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)  
 C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s  
 C:Genetics:  
 A:Gene: bont/b  
 C:Superfamily: tetanus toxin  
 C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc  
 F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>  
 F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HYV>  
 F:230,234/Binding site: zinc (His) #status predicted  
 F:231/Active site: Glu #status predicted

Query Match 61.8%; Score 34; DB 2; Length 1291;  
 Best Local Similarity 70.0%; Pred. No. 76;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 KFKFKFKFKF 10  
 I I I I I I I  
 Db 321 KNKFKDKYKF 330  
 RESULT 6  
 D64523  
 conserved hypothetical secreted protein HP0028 - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 31-Oct-1997  
 C:Accession: D64523  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: D64523  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-177 <TOM>  
 A:Cross-references: GB:AE000511; NID:g2313102; PID:g2313105; TIGR:HP0028  
 C:Superfamily: conserved hypothetical protein HI0453

Query Match 61.8%; Score 34; DB 2; Length 177;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 KFKFKFKFKF 10  
 I I I I I I I  
 Db 321 KNKFKDKYKF 330  
 RESULT 6  
 D64523  
 conserved hypothetical secreted protein HP0028 - Helicobacter pylori (strain 26695)  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 31-Oct-1997  
 C:Accession: D64523  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: D64523  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-177 <TOM>  
 A:Cross-references: GB:AE000511; NID:g2313102; PID:g2313105; TIGR:HP0028  
 C:Superfamily: conserved hypothetical protein HI0453

Db 88 RFSFPFYKFK 97  
 I I I I I I I  
 RESULT 7  
 A71153  
 hypothetical protein PH0424 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 15-Jan-1999  
 C:Accession: A71153  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamanoto, S.; Se  
 M.; Onfuku, T.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: A71153  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-348 <RAW>  
 A:Cross-references: GB:AP000002; NID:g3236129; PID:d1030453; PID:g3256827  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBa  
 C:Genetics:  
 A:Gene: PH0424  
 C:Superfamily: Methanobacterium thermoautotrophicum rhamnosyl transferase  
 Query Match 61.8%; Score 34; DB 2; Length 348;  
 Best Local Similarity 77.8%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 KFKFKFKFKF 9  
 I I I I I I I  
 Db 287 KSKEFKFYK 295  
 RESULT 8  
 A25157  
 macrolide/lincosamide/streptogramin B resistance protein - Bacteroides fragilis  
 C:Species: Bacteroides fragilis  
 C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 14-Nov-1997  
 C:Accession: A25157  
 R:Rasmussen, J.L.; Odelson, D.A.; Macrina, F.L.  
 J. Bacteriol. 168, 523-533, 1986  
 A:Title: Complete nucleotide sequence and transcription of ermF, a macrolide-lincosam  
 A:Reference number: A25157; MUID:87056929  
 A:Accession: A25157  
 A:Molecule type: DNA  
 A:Residues: 1-266 <RAS>  
 C:Genetics:  
 A:Gene: ermF  
 C:Superfamily: rRNA (adenine-N6-)-methyltransferase  
 C:Keywords: antibiotic resistance  
 Query Match 60.0%; Score 33; DB 2; Length 266;  
 Best Local Similarity 66.7%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 FKFKFKFKF 10  
 I I I I I I I  
 Db 189 FDFKFKAKY 197  
 RESULT 9  
 S55397  
 macrolide/lincosamide/streptogramin B resistance protein - Bacteroides sp. plasmid pi  
 C:Species: Bacteroides sp. plasmid pIP417  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 14-Nov-1997  
 C:Accession: S55397  
 R:Haggoud, A.; Trinh, S.; Mohieddine, M.; Reysset, G.  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Genetic analysis of the minimal replicon of plasmid pIP417 and compari

A:Reference number: S55397

A:Accession: S55397

A:Molecule type: DNA

A:Residues: 1-266 <HAG>

A:Cross-references: EMBL:X87253; NID:g853781; PID:g853782

C:Genetics:

A:Gene: ermF

A:Genome: plasmid

C:Superfamily: rRNA (adenine-N6-)-methyltransferase

C:Keywords: antibiotic resistance

Query Match 60.0%; Score 33; DB 2; Length 266;

Best Local Similarity 66.7%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKF 10

| |||| |

Db 189 FDFKFKAKY 197

RESULT 10

S57085

1-phosphatidylinositol 3-kinase (EC 2.7.1.137) TOR1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein J1803; protein YJR066W

C:Species: Saccharomyces cerevisiae

C:Date: 13-Sep-1995 #sequence\_revision 01-Mar-1996 #text\_change 24-Sep-1998

C:Accession: S57085; A54428; S43940; S71688

R:Manus. V.; Huang, M.E.; Galibert, F.

submitted to the Protein Sequence Database, September 1995

A:Reference number: S57085

A:Accession: S57085

A:Molecule type: DNA

A:Residues: 1-2470 <MAN>

A:Cross-references: EMBL:Z49566; NID:g1015742; PID:g1015743; MIPS:YJR066W

R:Carfkey, R.; Young, P.R.; McLaughlin, M.M.; Bergsma, D.J.; Koltin, Y.; Sathe, G.M.;

Mol. Cell. Biol. 13, 6012-6023, 1993

A:Title: Dominant missense mutations in a novel yeast protein related to mammalian phosph

A:Reference number: A54428; MUID:94019276

A:Accession: A54428

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-57,'G',59-114,'I',116-132,'N',134-395,'K',397-546,'S',548-1467,'R',1469-16

A:Cross-references: GB:L19540; NID:g408955; PID:g408956

A>Note: the authors translated the codon CGG for residue 1468 as Ala

R:Helliwell, S.B.; Wagner, P.; Kunz, J.; Deuter-Reinhard, M.; Henriquez, R.; Hall, M.N.

Mol. Biol. Cell 5, 105-118, 1994

A:Title: TOR1 and TOR2 are structurally and functionally similar but not identical phosph

A:Reference number: S43940; MUID:94243030

A:Accession: S43940

A:Molecule type: DNA

A:Residues: 1-230,'R',232-395,'K',397-546,'S',548-674,'I',676-1291,'E',1293-1435,'A',143

A:Cross-references: EMBL:X74857; NID:g468738; PID:g468739

R:Huang, M.E.; Manus, V.; Chuat, J.C.; Galibert, F.

Yeast 12, 869-875, 1996

A:Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames

A:Reference number: S71676; MUID:96437976

A:Accession: S71688

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2470 <HUA>

A:Cross-references: EMBL:147993; NID:g1019675; PID:g1019688

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: SGD:TOR1; DRR1

A:Cross-references: SGD:S0003827; MIPS:YJR066W

A:Map position: 10R

C:Function:

A:Description: required for translation initiation; required for G1 progression

C:Superfamily: yeast TOR2 protein

C:Keywords: DNA binding; EF hand; leucine zipper; phosphotransferase

Query Match 60.0%; Score 33; DB 2; Length 2470;

Best Local Similarity 66.7%; Pred. No. 2.1e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKFKFKF 9

|||| |

Db 441 KFKFKFKE 449

RESULT 11

I40610

transcription initiation factor sigma A - Clostridium acetobutylicum

N:Alternate names: major vegetative sigma factor

C:Species: Clostridium acetobutylicum

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 10-Oct-1997

C:Accession: I40610; S34307

R:Sauer, U.; Treuner, A.; Buchholz, M.; Santangelo, J.D.; Durre, P.

J. Bacteriol. 176, 6572-6582, 1994

A:Title: Sporulation and primary sigma factor homologous genes in Clostridium acetobu

A:Reference number: I40609; MUID:95050216

A:Accession: I40610

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-378 <RES>

A:Cross-references: EMBL:Z23080; NID:g312371; PID:g312373

C:Genetics:

A:Gene: sigma

C:Superfamily: transcription initiation factor sigma 43; transcription initiation fac

C:Keywords: DNA binding; sigma factor; transcription initiation

F:20-146/Domain: transcription initiation factor sigma region 1 homology <SRI>

F:147-373/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match

Best Local Similarity 50.0%; Score 33; DB 2; Length 378;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKF 10

|||| |

Db 184 KDFKKGFKF 193

RESULT 12

S36569

E1 protein - human papillomavirus type 49

C:Species: human papillomavirus type 49

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Sep-1997

C:Accession: S36569

R:Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36569

A:Molecule type: DNA

A:Residues: 1-609 <DBL>

A:Cross-references: EMBL:X74480; NID:g397030; PID:g397033

C:Superfamily: papillomavirus E1 protein

C:Keywords: early protein; nucleus

Query Match

Best Local Similarity 60.0%; Score 33; DB 2; Length 609;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10

|||| |

Db 546 FEFKFKFPE 554

RESULT 13

S62919

probable membrane protein YNL008c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein N2874  
C:Species: Saccharomyces cerevisiae  
C:Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 07-Aug-1998  
C:Accession: S62919; S62920  
R:Dolignon, F.; Crouzet, M.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62916  
A:Accession: S62919  
A:Molecule type: DNA  
A:Residues: 1-370 <DOI>  
A:Cross-references: EMBL:Z71284; MIPS:YNL008C  
R:Andre, B.; Iraqi Houssein, I.; Urrestarazu, L.A.; Visser, S.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62920  
A:Accession: S62920  
A:Molecule type: DNA  
A:Residues: 103-669 <AND>  
A:Cross-references: EMBL:Z71284; MIPS:YNL008C  
A:Experimental source: strain S288C  
C:Genetics:  
A:Map position: 14L  
C:Superfamily: probable membrane protein YMR119W  
C:Keywords: transmembrane protein  
F:82-98/Domain: transmembrane #status predicted <TM1>  
F:117-133/Domain: transmembrane #status predicted <TM2>  
F:153-169/Domain: transmembrane #status predicted <TM3>  
F:211-227/Domain: transmembrane #status predicted <TM4>  
F:275-291/Domain: transmembrane #status predicted <TM5>

Query Match 60.0%; Score 33; DB 2; Length 669;  
Best Local Similarity 60.0%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KFKFKFKF 10  
| :||| | |  
Db 595 KLEKFDFAF 604

RESULT 14  
A70169  
Hypothetical protein BB0554 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: A70169  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: A70169  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-626 <KLE>  
A:Cross-references: GB:AE001157; GB:AE000783; NID:g2688471; PID:g2688481; TIGR:BB0554  
A:Experimental source: strain B31

Query Match 60.0%; Score 33; DB 2; Length 626;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FKFKFKFK 10  
| :||| | |  
Db 262 FLYKFKFK 270

RESULT 15  
A45031  
cysteine-rich fibroblast growth factor receptor - chicken

C:Species: Gallus gallus (chicken)  
C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
C:Accession: A45031  
R:Burrus, L.W.; Zuber, M.E.; Lueddecke, B.A.; Olwin, B.B.  
Mol. Cell. Biol. 12, 5600-5609, 1992  
A:Title: Identification of a cysteine-rich receptor for fibroblast growth factors.  
A:Reference number: A45031; MUID:93078761  
A:Accession: A45031  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-1142 <BUR>  
A:Cross-references: GB:M95766; NID:g211775; PID:g211776  
A:Experimental source: embryos, brain  
A:Note: sequence extracted from NCBI backbone (NCBIN:119082, NCBIP:119083)  
C:Keywords: growth factor receptor

Query Match 60.0%; Score 33; DB 2; Length 1142;  
Best Local Similarity 62.5%; Pred. No. 1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FKFKFKFK 9  
| :||| | |  
Db 712 FRFSYKFK 719

Search completed: September 7, 1999, 23:06:19  
Job time: 2467 sec

**THIS PAGE BLANK (USPTO)**

```
Query Match      74.5%; Score 41; DB 1; Length 436;
Best Local Similarity 77.8%; Pred. No. 0.88;
Matches 7: Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY 2 FKFKFKFK 10  
1 :|||||  
Db 209 FPKFKFK 217

RESULT 2  
YNM8\_YEAST  
ID YNM8\_YEAST STANDARD; PRT; 434 AA.  
AC PS3916;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 50.2 KD PROTEIN IN CPT1-SPC98 INTERGENIC REGION.  
GN YNL128W OR N1220 OR N1872  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C.  
RX MEDLINE; 96109932.  
RA MALLEY L., BUSSEREAU F., JACQUET M.;  
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
RT MEP2, CAP/SRV2, NAM9, FKBI/FPRI/RBPI, MOM22 and CPT1, predicts an  
RT adenosine deaminase gene and 14 new open reading frames.";  
RL YEAST 11:1195-1209(1995).  
CC -!- SIMILARITY: SOME, TO PROTEIN-TYROSINE PHOSPHATASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 246843; G854509; -  
DR EMBL; 271404; E239804; -  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
KW HYPOTHETICAL PROTEIN; HYDROLASE.  
SQ SEQUENCE 434 AA; 50152 MW; 368BEB97 CRC32;

Query Match 63.6%; Score 35; DB 1; Length 434;  
Best Local Similarity 85.7%; Pred. No. 9.8;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FKFKFK 7  
|||||  
Db 268 FKFKFK 274

RESULT 3  
BXB\_CLOBO  
ID BXB\_CLOBO STANDARD; PRT; 1290 AA.  
AC P10844; P10843;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE BOTULINUM NEUROTOXIN TYPE B PRECURSOR (EC 3.4.24.69) (BONT/B)  
DE (BONTOLYLXIN B).  
GN BOTB  
OS CLOSTRIDIUM BOTULINUM.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;  
OC CLOSTRIDIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92384550  
RA WHELAN S.M., ELMORE M.J., BODSWORTH N.J., BREHM J.K., ATKINSON T.,  
RA MINTON N.P.;  
RT "Molecular cloning of the Clostridium botulinum structural gene  
encoding the type B neurotoxin and determination of its entire  
nucleotide sequence.";  
APPL. ENVIRON. MICROBIOL. 58:2345-2354(1992).  
[2]  
RP SEQUENCE OF 35-245 FROM N.A.  
RC STRAIN=NCTC 7273;  
RA SZABO E.A., PEMBERTON J.M., DESMARCHELIER P.M.;  
RL SUBMITTED (APR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
[3]  
RP SEQUENCE OF 633-993 FROM N.A.  
RC STRAIN=NCTC 7273;  
RX CAMPBELL K., EAST A.K., COLLINS M.D.;  
RA "Gene probes for identification of the botulinum neurotoxin gene and  
RT specific identification of neurotoxin types B, E, and F.";  
RL J. CLIN. MICROBIOL. 31:2255-2262(1993).  
[4]  
RP SEQUENCE OF 1-44 AND 441-458.  
RC STRAIN=657;  
RX MEDLINE; 89000987.  
RA DASGUPTA B.R., DATTA A.;  
RT "Botulinum neurotoxin type B (strain 657): partial sequence and  
RT similarity with tetanus toxin.";  
RL BIOCHIMIE 70:811-817(1988).  
[5]  
RP SEQUENCE OF 1-16 AND 441-458.  
RC STRAIN=OKRA;  
RX MEDLINE; 85197963.  
RA SCHMIDT J.J., SATHYAMOORTHY V., DASGUPTA B.R.;  
RT "Partial amino acid sequences of botulinum neurotoxins types B and  
RT E.";  
RL ARCH. BIOCHEM. BIOPHYS. 238:544-548(1985).  
[6]  
RP IDENTIFICATION AS ZINC-PROTEASE.  
RX MEDLINE; 93054694.  
RA SCHIAVO G., ROSSETTO O., SANTUCCI A., DASGUPTA B.R., MONTECUCCO C.;  
RT "Botulinum neurotoxins are zinc proteins.";  
RL J. BIOL. CHEM. 267:23479-23483(1992).  
[7]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE; 93063293.  
RA SCHIAVO G., BENFENATI F., POULAIN B., ROSSETTO O., DE LAURETO P.P.,  
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
RT by proteolytic cleavage of synaptobrevin.";  
RL NATURE 359:832-835(1992).  
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF  
CC SYNAPTOSOMAL-2.  
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A  
CC A HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL  
CC ACTIVITY, WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE  
CC CHANNEL FORMATION AND TOXIN BINDING, RESPECTIVELY.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF BOTULINUM  
CC NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27 (ZINC  
CC METALLOPROTEASE); ALSO KNOWN AS THE TETANUS/BOTULINUM NEUROTOXIN  
CC SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M81186; G144735; -



```
DR EMBL; Z11934; G40384; -.
DR EMBL; X70817; G407783; -.
DR PIR; S07128; S07128.
DR PIR; S07155; S07155.
DR PIR; S08562; S08562.
DR PIR; S08573; S08573.
DR PIR; S08574; S08574.
DR PIR; S08574; S08574.
DR PIR; A48940; A48940.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PFAM; PF00099; zn-protease; 1.
DR HSSP; P04958; 1AP9.
KW NEUROTOXIN; TRANSMEMBRANE; HYDROLASE; METALLOPROTEASE; ZINC.
FT INIT_MET 0
FT CHAIN 1 440
FT CHAIN 441 1290
FT METAL 229 229
FT ACT_SITE 230 230
FT METAL 233 233
FT METAL 436 445
FT DISULFID 29 29
FT CONFLICT 217 217
FT CONFLICT 224 224
FT CONFLICT 463 463
SQ SEQUENCE 1290 AA; 150670 MW; 479BBC70 CRC32;

Query Match 61.8%; Score 34; DB 1; Length 1290;
Best Local Similarity 70.0%; Pred. No. 42;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFKFKFKF 10
Db 320 KNKFKDYKF 329

RESULT 4
EXON_HSV6U
ID EXON_HSV6U STANDARD; PRT; 488 AA.
AC P24447;
DT 01-MAR-1992 (REL. 21, LAST CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ALKALINE EXONUCLEASE (EC 3.1.11.-).
GN U70 OR 16R.
OS HERPES SIMPLEX VIRUS (TYPE 6 / STRAIN UGANDA-1102).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC BETAHERPESVIRINAE; ROSEOLVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90080132.
RA LAWRENCE G.L., CHEE M., CRAXTON M.A., GOMPELS U.A., HONESS R.W.,
RA BARRELL B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";
RL J. VIROL. 64:287-299(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95266321.
RA GOMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,
RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
and genome evolution.";
RL VIROLOGY 209:29-51(1995).
CC !- SIMILARITY: TO OTHER HERPESVIRUSES ALKALINE EXONUCLEASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

RN SEQUENCE FROM N.A.  
 RP STRAIN=M;  
 RX MEDLINE; 95050273.  
 RA LIN W.S., CUNNEEN T., LEE C.Y.;  
 RT "sequence analysis and molecular characterization of genes required  
 for the biosynthesis of type 1 capsular polysaccharide in  
 Staphylococcus aureus".  
 RL J. BACTERIOL. 176:7005-7016(1994).  
 CC -!- FUNCTION: REQUIRED FOR THE BIOSYNTHESIS OF TYPE 1 CAPSULAR  
 POLYSACCHARIDE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U10927; G506707; -;  
 SQ SEQUENCE 449 AA; 52510 MW; 1B27A9C4 CRC32;  
 -----  
 Query Match 60.0%; Score 33; DB 1; Length 449;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 -----  
 QY 1 KFKFKFKFKF 10  
 Db 410 KLYELKFRF 419  
 | : : : : |  
: : : :
 RESULT 7  
 ERMF\_BACFR  
 ID ERMF\_BACFR STANDARD; PRT; 266 AA.  
 AC P10337;  
 DT 01-MAR-1989 (REL. 10, CREATED)  
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE RNA ADENINE N-6-METHYLTRANSFERASE (EC 2.1.1.48) (MACROLIDE-  
 DE LINCOSAMIDE-STREPTOGRAMIN B RESISTANCE PROTEIN) (ERYTHROMYCIN  
 DE RESISTANCE PROTEIN).  
 GN ERMF  
 OS BACTEROIDES FRAGILIS.  
 OG PLASMID PB4.  
 OC BACTERIA; CYTOPHAGALES; BACTEROIDACEAE; BACTEROIDES.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP RASMUSSEN J.L., ODELMAN D.A., MACRINA F.L.;  
 RX MEDLINE; 87056929.  
 RA "Complete nucleotide sequence and transcription of ermF, a macrolide-  
 RT lincosamide-streptogramin B resistance determinant from Bacteroides  
 RT fragilis".  
 RL J. BACTERIOL. 168:523-533(1986).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TRANSPOSON-TN4351;  
 RX MEDLINE; 87279905.  
 RA RASMUSSEN J.L., ODELMAN D.A., MACRINA F.L.;  
 RT "Complete nucleotide sequence of insertion element IS4351 from  
 RT Bacteroides fragilis".  
 RL J. BACTERIOL. 169:3573-3580(1987).  
 CC -!- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE  
 CC RESIDUE AT POSITION 2058 IN 23S RNA, RESULTING IN REDUCED  
 CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B  
 CC ANTIBIOTICS.  
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + RNA -> S-ADENOSYL-L-  
 CC HOMOCYSTEINE + RNA CONTAINING N(6)-METHYLADENINE.  
 CC -!- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE  
 CC FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M14730; G150541; -;  
 DR EMBL; M17124; G154890; -;  
 DR PIR; A25157; A25157.  
 DR PROSITE; PS01131; RNA\_A\_DIMETH; 1.  
 DR PFAM; PF00398; RnaAD; 1.  
 KW ANTIBIOTIC RESISTANCE; TRANSFERASE; METHYLTRANSFERASE; PLASMID;  
 KW TRANSPOSABLE ELEMENT.  
 SQ SEQUENCE 266 AA; 30356 MW; EB272760 CRC32;  
 -----  
 Query Match 60.0%; Score 33; DB 1; Length 266;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 -----  
 QY 2 FKFKEFKF 10  
 Db 189 FDFKFKAKY 197  
 | : : : : |  
: : : :
 RESULT 8  
 RPSD\_CLOAB  
 ID RPSD\_CLOAB STANDARD; PRT; 378 AA.  
 AC P33656;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-A) (MAJOR VEGETATIVE SIGMA  
 DE FACTOR).  
 GN RPOD OR SIGA.  
 OS CLOSTRIDIUM ACETOBUTYLICUM.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;  
 OC CLOSTRIDIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 792;  
 RX MEDLINE; 95050216.  
 RA SAUER U., TREUNER A., BUCHHOLZ M., SANTANGELO J.D., DURR P.;  
 RT "Sporulation and primary sigma factor homologous genes in Clostridium  
 RT acetobutylicum".  
 RL J. BACTERIOL. 176:6572-6582(1994).  
 CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES  
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND  
 CC THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS  
 CC BACTERIA.  
 CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z23080; G312373; -;  
 DR PIR; S34307; S34307.  
 DR PROSITE; PS00715; SIGMA70\_1; 1.  
 DR PROSITE; PS00716; SIGMA70\_2; 1.  
 DR PFAM; PF00140; sigma70; 1.  
 DR HSP; P00579; 1S1G.  
 KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;  
 KW DNA-BINDING. 169 182 POLYMERASE CORE BINDING (POTENTIAL).  
 FT DOMAIN 339 358 H-T-H MOTIF (BY SIMILARITY).  
 SQ DNA\_BIND 378 AA; 43260 MW; C438C6F7 CRC32;  
 FT SEQUENCE

Query Match 60.0%; Score 33; DB 1; Length 378;  
 Best Local Similarity 70.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKPKFKFK 10  
 DB 184 KFDKKGKFK 193

# RESULT 9

TORI\_YEAST STANDARD; PRT; 2470 AA.  
 AC P35169;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE PHOSPHATIDYLINOSITOL 3-KINASE TORI (EC 2.7.1.137) (PI3-KINASE)  
 DE (PTDINS-3-KINASE) (PI3K).  
 GN TORI OR DRR1 OR YJRO66W OR J1803.  
 OS SACCHAROMYCES CREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 94019276.  
 RA CAFFERKEY R., YOUNG P.R., MCLAUGHLIN M.M., BERGSMAN D.J., KOLTIN Y.,  
 RA SATHE G.M., FAUCETTE L., ENG W.-K., JOHNSON R.K., LIVI G.P.;  
 RT "Dominant missense mutations in a novel yeast protein related to  
 mammalian phosphatidylinositol 3-kinase and vps34 abrogate rapamycin  
 cytotoxicity.";  
 RL MOL. CELL. BIOL. 13:6012-6023(1993).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX STRAIN-JK-3D;  
 RX MEDLINE; 94243030.  
 RA HELLWELL S.B., WAGNER P., KUNZ J., DEUTER-REINHARD M., HENRIQUEZ R.,  
 RA HALL M.N.;  
 RT "TOR1 and TOR2 are structurally and functionally similar but not  
 identical phosphatidylinositol kinase homologues in yeast.";  
 RL MOL. BIOL. CELL 5:105-118(1994).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX STRAIN-S288C;  
 RX MEDLINE; 96437976.  
 RA HUANG M.-E., MANUS V., CHUAT J.-C., GALIBERT F.;  
 RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open  
 reading frames and a gene cluster with a counterpart on chromosome  
 XI.";  
 RL YEAST 12:869-875(1996).  
 CC -1- FUNCTION: PHOSPHATIDYLINOSITOL 3-KINASE HOMOLOG REQUIRED FOR G1  
 CC PROGRESSION: TARGET OF THE ANTIBIOTIC RAPAMYCIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + 1-PHOSPHATIDYL-D-MYO-INOSITOL = ADP +  
 CC 1-PHOSPHATIDYL-D-ADP + 1-PHOSPHATIDYL-D-3-PHOSPHATE.  
 CC -1- IT MAY ACT ON ANOTHER SUBSTRATE OR PHOSPHORYLATE A DIFFERENT  
 CC POSITION IN THE PHOSPHATIDYLINOSITOL RING.  
 CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.  
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MEK-1 IS THE INITIATOR.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; L19540; G408956; -;  
 DR EMBL; X74857; G488739; -;  
 DR EMBL; 249566; G1015743; -;  
 DR EMBL; L47993; G1019688; -;  
 DR SGD; L0002322; TOR1.

DR PROSITE; PS00915; PI3\_4\_KINASE\_1; 1.  
 DR PROSITE; PS00916; PI3\_4\_KINASE\_2; 1.  
 DR PFAM; PF00454; PI3\_PI4\_Kinase; 1.  
 DR HSP; P42345; 1FAP.  
 KW TRANSFERASE; KINASE; CELL CYCLE.  
 FT DOMAIN 441 447 ARG/LYS-RICH (BASIC).  
 FT DOMAIN 2125 2470 PI3K/PI4K.  
 FT CONFLICT 58 58 S->R; LOSS OF RAPAMYCIN SENSITIVITY.  
 FT CONFLICT 115 115 D->G (IN REF. 1).  
 FT CONFLICT 133 133 V->I (IN REF. 1).  
 FT CONFLICT 231 231 S->N (IN REF. 1).  
 FT CONFLICT 396 396 A->R (IN REF. 2).  
 FT CONFLICT 547 547 N->K (IN REF. 1 AND 2).  
 FT CONFLICT 675 675 T->S (IN REF. 1 AND 2).  
 FT CONFLICT 1292 1292 G->E (IN REF. 2).  
 FT CONFLICT 1436 1436 G->A (IN REF. 2).  
 FT CONFLICT 1468 1468 A->R (IN REF. 1 AND 2).  
 FT CONFLICT 1469 1471 WGL->GGG (IN REF. 2).  
 FT CONFLICT 1478 1479 EQ->DE (IN REF. 2).  
 FT CONFLICT 1590 1590 V->I (IN REF. 2).  
 FT CONFLICT 1642 1642 NDPSLPNTERA->TILVQIRSKP (IN REF. 2).  
 FT CONFLICT 1640 1640 F->V (IN REF. 1).  
 FT CONFLICT 1844 1844 L->S (IN REF. 2).  
 FT CONFLICT 1972 1972 S->R (IN REF. 1).  
 FT CONFLICT 2202 2202 H->Q (IN REF. 1).  
 FT CONFLICT 2414 2414 K->R (IN REF. 1 AND 2).  
 SQ SEQUENCE 2470 AA; 281136 MW; F48E060A CRC32;  
 Query Match 60.0%; Score 33; DB 1; Length 2470;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KFKPKFKFK 9  
 DB 441 KFKPKFKFK 449  
 RESULT 10  
 VEL\_HP49 STANDARD; PRT; 609 AA.  
 ID VEL\_HP49  
 AC P36729;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE REPLICATION PROTEIN EI.  
 GN EI.  
 OS HUMAN PAPILLOMAVIRUS TYPE 49.  
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; PAPAPOVIRIDAE; PAPILLOMAVIRUS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 94265501.  
 RA DELIUS H., HOFMANN B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RL CURR. TOP. MICROBIOL. IMMUNOL. 186:13-31(1994).  
 CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF  
 CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2  
 CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH  
 CC CONTAINS BINDING SITES FOR BOTH PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X74480; G397033; -;  
 DR PIR; S36569; S36569.  
 DR PFAM; PF00519; EI; 1.

DR PFAM; PF00524; EL\_N; 1.  
 KW EARLY PROTEIN; DNA REPLICATION; HELICASE; ATP-BINDING; DNA-BINDING;  
 KW NUCLEAR PROTEIN.  
 FT NP\_BIND 437 444 ATP (POTENTIAL).  
 SQ SEQUENCE 609 AA; 70146 MW; 4E2BBACE CRC32;

Query Match 60.0%; Score 33; DB 1; Length 609;  
 Best Local Similarity 66.7%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FKFKFKFK 10  
 I:|||||  
 Db 546 FEKHKFKF 554

## RESULT 11

YNAB YEAST  
 ID YNAB YEAST STANDARD; PRT; 669 AA.  
 AC P53983;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 76.7 KD PROTEIN IN SPO1-SIS1 INTERGENIC REGION.  
 GN YN008C OR N2874.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 [1]  
 RN SEQUENCE OF 1-102 FROM N.A.  
 RP ANDRE B.; IRAQI HOUSSEINI I.; URRESTARAZU L.A.; VISSERS S.;  
 RA SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RL  
 RN [2]  
 RP SEQUENCE OF 371-669 FROM N.A.  
 RA DOIGNON F.; CROUZET M.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; 271284; E239659; -;  
 CC EMBL; 271283; E328821; -;  
 CC KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 72 92 POTENTIAL.  
 FT TRANSMEM 120 140 POTENTIAL.  
 FT TRANSMEM 150 170 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 209 229 POTENTIAL.  
 FT TRANSMEM 234 254 POTENTIAL.  
 FT TRANSMEM 271 291 POTENTIAL.  
 SQ SEQUENCE 669 AA; 76741 MW; 1E1F91BE2 CRC32;

Query Match 60.0%; Score 33; DB 1; Length 669;  
 Best Local Similarity 60.0%; Pred. No. 33;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
 I:|||||  
 Db 595 KLEKDFAF 604

## RESULT 12

PP16 ARATH  
 ID PP16 ARATH STANDARD; PRT; 322 AA.  
 AC P48486;  
 DT 01-FEB-1996 (REL. 33, CREATED)

DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE SERINE/THREONINE PROTEIN PHOSPHATASE PP1 ISOZYME 6 (EC 3.1.3.16).  
 GN PPIBG  
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRASS).  
 OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 95291339.  
 RA ARUNDHATI A.; FEILER H.; TRAAS J.; ZHANG H.; LUNNESS P.A.;  
 RA DOONAN J.H.;  
 RT "A novel Arabidopsis type 1 protein phosphatase is highly expressed  
 RT in male and female tissues and functionally complements a conditional  
 RT cell cycle mutant of *Aspergillus*.";  
 RL PLANT J. 7:823-834(1995).  
 CC -1- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H<sub>2</sub>O -> A PROTEIN +  
 CC ORTHOPHOSPHATE (THIS ENZYME IS SERINE/THREONINE SPECIFIC).  
 CC -1- TISSUE SPECIFICITY: STRONGLY UPREGULATED WITHIN DEVELOPING  
 CC FLOWERS, ESPECIALLY IN THE TAPETUM, THE DEVELOPING AND MATURE  
 CC POLLEN AND IN THE OVARIES.  
 CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-1  
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; 246253; G829255; -;  
 CC PROSITE; PS00125; SER\_THR\_PHOSPHATASE; 1.  
 CC PFAM; PF00149; Stphosphatase; 1.  
 CC DR HSP; P08129; 1FJM.

KW HYDROLASE; IRON; MANGANESE; MULTIGENE FAMILY.  
 FT METAL 61 61 IRON (BY SIMILARITY).  
 FT METAL 63 63 IRON (BY SIMILARITY).  
 FT METAL 89 89 IRON AND MANGANESE (BY SIMILARITY).  
 FT METAL 121 121 MANGANESE (BY SIMILARITY).  
 FT ACT\_SITE 122 122 GENERAL ACID (BY SIMILARITY).  
 FT METAL 170 170 MANGANESE (BY SIMILARITY).  
 FT METAL 245 245 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 322 AA; 36582 MW; 1EE6F954 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 322;  
 Best Local Similarity 55.6%; Pred. No. 24;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FKFKFKFK 10  
 I:|||||  
 Db 107 YKVKYKFN 115

## RESULT 13

PRIA BORBU  
 ID PRIA BORBU STANDARD; PRT; 660 AA.  
 AC Q45032; OS1047;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PRIMOSOMAL PROTEIN N' (REPLICATION FACTOR V).  
 GN PRIA OR BB0014.  
 OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).  
 OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-HB19;  
 RA BOURSAX-UEDE C.;  
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

```

RN  [2]
RC  SEQUENCE FROM N.A.
RX  STRAIN-ATCC 35210 / B31;
RA  MEDLINE; 98065943.
RA  FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA  LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA  DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA  PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA  VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA  UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA  GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA  SMITH H.O., VENTER J.C.;
RT  "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT  burgdorferi.";
RL  NATURE 390:580-586(1997).
RC  SEQUENCE OF 165-614 FROM N.A.
RX  STRAIN-297;
RA  PAN M., YEH J., TSAI C.;
RL  SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC  -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC  ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC  THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC  BACTERIA.
CC  -1- SIMILARITY: BELONGS TO THE P1A SUBFAMILY OF HELICASES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X97449; E238797;
DR  EMBL; AE001115; G2687882;
DR  TIGR; BB0014;
DR  PFAM; PF00271; helicase_C; 1.
KW  DNA REPLICATION; DNA-BINDING; ATP-BINDING; HELICASE; PRIMOSOME;
FT  NP_BIND 158 165 ATP (POTENTIAL).
FT  SITE 256 259 DEH BOX.
FT  2N_FING 370 382 C4-TYPE (POTENTIAL).
FT  2N_FING 397 413 C4-TYPE (POTENTIAL).
FT  CONFLICT 126 126 P -> L (IN REF. 1).
FT  CONFLICT 555 555 D -> N (IN REF. 1).
SQ  SEQUENCE 660 AA; 77551 MW; B79DC6E CRC32;

Query Match 58.2%; Score 32; DB 1; Length 660;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKFKFKFK 9
Db 58 EFKEFKFK 66

RESULT 14
RPSD_BORBU
ID RPSD_BORBU STANDARD; PRT; 631 AA.
AC P52323;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-70).
GN RPOD OR BB0712.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN [1]
RP SEQUENCE OF 89-631 FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RA PAN M.;
RL THESIS (1994), NATIONAL TAIWAN UNIVERSITY, TAIWAN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;

```

```

RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL NATURE 390:580-586(1997).
RC SEQUENCE OF 165-614 FROM N.A.
RX STRAIN-297;
RA PAN M., YEH J., TSAI C.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC BACTERIA.
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U17591; G596248;
DR  EMBL; AE001171; G2688647;
DR  EMBL; U68006; G1553017;
DR  TIGR; BB0712;
DR  PROSITE; PS00715; SIGMA70_1; 1.
DR  PROSITE; PS00716; SIGMA70_2; 1.
DR  PFAM; PF00140; sigma70; 1.
DR  HSP; P00579; ISIG.
KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
FT DOMAIN 419 432 POLYMERASE CORE BINDING (POTENTIAL).
FT DNA_BIND 589 608 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 631 AA; 73642 MW; B51AF446 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 631;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFKFKFKFK 10
Db 434 KFEYKRGFK 443

RESULT 15
RT07_ACACA
ID RT07_ACACA STANDARD; PRT; 337 AA.
AC P46756;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.
GN RPS7.
OS ACANTHAMOEBA CASTELLANII (AMOEB).
OC MITOCHONDRION.
OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 30010 / NEFF;
RX MEDLINE; 93157849.
RA LONERGAN K.M., GRAY M.W.;
RT "Editing of transfer RNAs in Acanthamoeba castellanii mitochondria.";

```

RL SCIENCE 259:812-816(1993).  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
CC -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U12386; G562050; -;  
DR PROSITE: PS00052; RIBOSOMAL\_S7; FALSE\_NEG.  
DR PIRAM: PF00177; S7; 1.  
KW RIBOSOMAL PROTEIN; MITOCHONDRION.  
SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 337;  
Best Local Similarity 75.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KEFKRKEF 8  
Db 229 KFKYRKEF 236  
|||||

Search completed: September 7, 1999, 23:50:06  
Job time: 1945 sec

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RN NATURE 390:580-586(1997).
RP SEQUENCE OF 165-614 FROM N.A.
RC STRAIN-297;
RA PAN M., YEH J., TSAI C.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC PRIB AND PRIC. FORMATION OF DNAB, DNAC, DNAT AND PRIMASE. PRTA THEN
CC SUBSEQUENT ACTIONS OF DNAB, DNAC, DNAT AND PRIMASE. PRTA THEN
CC FUNCTIONS AS A HELICASE WITHIN THE PRIMOSOME (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PRTA SUBFAMILY OF HELICASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; X97449; F238797;
DR EMBL; AE001115; G2687882;
DR TIGR; BF0014;
DR PFAM; PF00271; helicase_C; 1.
KW DNA REPLICATION; DNA-BINDING; ATP-BINDING; HELICASE; PRIMOSOME;
KW ZINC-FINGER.
FT NP_BIND 158 165 ATP (POTENTIAL).
FT SITE 256 259 DEER BOX.
FT 2N_FING 370 382 C4-TYPE (POTENTIAL).
FT 2N_FING 397 413 C4-TYPE (POTENTIAL).
FT CONFLICT 126 126 P -> L (IN REF. 1).
FT CONFLICT 555 555 D -> N (IN REF. 1).
SQ SEQUENCE 660 AA; 77551 MW; B799DC6E CRC32;

Query Match 58.2%; Score 32; DB 1; Length 660;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKFKFKFK 9
Db 58 EFKEKFKFK 66
:::|::|::|

RESULT 14
RPSD_BORBU STANDARD; PRT; 631 AA.
AC P52323;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE RPA POLYMERASE SIGMA FACTOR RPOD (SIGMA-70).
GN RPOD OR BB0712.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN [1]
RP SEQUENCE OF 89-631 FROM N.A.
RC STRAIN-ATCC 35210 / B31;
RA PAN M.;
RL THESIS (1994), NATIONAL TAIWAN UNIVERSITY, TAIWAN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;

```

```

RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RN NATURE 390:580-586(1997).
RP SEQUENCE OF 165-614 FROM N.A.
RC STRAIN-297;
RA PAN M., YEH J., TSAI C.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC BACTERIA.
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; U17591; G596248;
DR EMBL; AE001171; G2688647;
DR EMBL; U68006; G1553017;
DR TIGR; BB0712;
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
DR PFAM; PF00140; sigma70; 1.
DR HSP; P00579; ISIG.
KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
FT DOMAIN 419 432 POLYMERASE CORE BINDING (POTENTIAL).
FT DNA_BIND 589 608 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 631 AA; 73642 MW; B51AF446 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 631;
Best Local Similarity 60.08; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFKFKFKFK 10
Db 434 KFEYKRGFKF 443
:::|::|::|

RESULT 15
RT07_ACACA STANDARD; PRT; 337 AA.
AC P46756;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.
GN RPS7.
OS ACANTHAMOEBA CASTELLANII (AMOEBA).
OC MITOCHONDRION.
OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 30010 / NEFF;
RX MEDLINE; 93157849.
RA LONERGAN K.M., GRAY M.W.;
RT "Editing of transfer RNAs in Acanthamoeba castellanii mitochondria.";

```

RL SCIENCE 259:812-816(1993).  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
CC -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U12386; G562050; -  
DR PROSITE; PS00052; RIBOSOMAL\_S7; FALSE\_NEG.  
DR PFAM; PF00177; S7; 1.  
KW RIBOSOMAL PROTEIN; MITOCHONDRION  
SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 337;  
Best Local Similarity 75.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KFKFKFKF 8  
Db 229 KFKYKRF 236  
|||:|  
|||:|

Search completed: September 7, 1999, 23:50:06  
Job time: 1945 sec



GenCore version 4.5

Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:32 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-5

Perfect score: 55

Sequence: 1 KFKFKFKFKF 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	74.5	275	2	033966 desulfovibr
2	41	74.5	208	2	085964 desulfovibr
3	41	74.5	212	2	085966 desulfovibr
4	41	74.5	223	2	085970 desulfovibr
5	41	74.5	367	2	085974 desulfovibr
6	41	74.5	213	2	085976 desulfovibr
7	41	74.5	221	2	085978 desulfovibr
8	41	74.5	471	2	092H18 desulfovibr
9	41	74.5	325	8	078874 desulfovibr
10	39	70.9	175	5	060943 desulfovibr
11	37	67.3	411	5	P90776 caenorhabd
12	37	67.3	399	8	Q34129 caluromys
13	37	67.3	399	8	Q34130 caluromys
14	37	67.3	238	10	082294 arabidopsis
15	35	63.6	397	1	033909 pyrobaculum
16	35	63.6	416	2	067601 aquifex
17	35	63.6	434	3	Q02933 saccharomyc
18	35	63.6	1768	5	Q24153 drosophila
19	35	63.6	1741	5	046095 drosophila
20	35	63.6	455	5	045353 caenorhabd
21	35	63.6	1321	5	097322 plasmodium
22	34	61.8	348	1	058161 pyrococcus
23	34	61.8	177	2	024871 helicobacte
24	34	61.8	1291	2	Q08077 clostridium
25	34	61.8	360	2	Q92FNO peptostrept
26	34	61.8	1291	2	092A78 clostridium
27	34	61.8	143	5	Q93340 caenorhabd
28	33	60.0	626	2	051504 borrelia bu
29	33	60.0	720	2	069273 corynebacte

30	33	60.0	266	2	Q56387	056387 escherichia
31	33	60.0	1179	4	Q92896	Q92896 homo sapien
32	33	60.0	1177	4	Q13221	Q13221 homo sapien
33	33	60.0	359	5	O62165	O62165 caenorhabd1
34	33	60.0	740	5	O20760	O20760 caenorhabd1
35	33	60.0	345	8	O78895	O78895 anolis bart
36	33	60.0	345	8	O78896	O78896 anolis verm
37	33	60.0	766	9	O80245	O80245 mycoplasma
38	33	60.0	301	10	O49476	O49476 arabidopsis
39	33	60.0	1171	11	O62638	O62638 rattus norv
40	33	60.0	1175	11	O61543	O61543 mus muscucu
41	33	60.0	1160	11	O92159	O92159 cricetus
42	33	60.0	609	12	O56936	O56936 human papil
43	33	60.0	609	12	O56942	O56942 human papil
44	33	60.0	1142	13	Q02391	Q02391 gallus gall
45	33	60.0	1077	13	Q91019	Q91019 gallus gall

  

RESULT 1						
O33966	PRELIMINARY;	PRT;	275	AA.		
AC	O33966					
DT	01-JAN-1998 (TremBLrel. 05, Created)					
DT	01-JAN-1998 (TremBLrel. 05, Last sequence update)					
DT	01-AUG-1998 (TremBLrel. 07, Last annotation update)					
DE	DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (EC 1.8.99.1)					
DE	(FRAGMENT).					
GN	DSRA.					
OS	Desulfovibrio simplex.					
OC	Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-XVI (DSM4141);					
RA	POTT W.M., THOM-SCHMITZ A.S., FAATH I., DAHL C., TRUEPER H.G.;					
RL	Microbiology 0:0-0(1997).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-XVI (DSM4141);					
RA	MOLITOR M., DAHL C., ZELLNER G., TRUEPER H.G.;					
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.					
CC	- - CATALYTIC ACTIVITY: H(2)S + ACCEPTOR + 3 H(2)O = SULFITE +					
CC	REDUCED ACCEPTOR.					
CC	- - COFACTOR: IRON.					
DR	EMBL; U78738; AB966716.1; -					
KW	Oxidoreductase.					
FT	NON_TER 275					
SQ	SEQUENCE 275 AA; 30612 MW; F7EC82B5 CRC32;					

  

Query Match	74.5%;	Score	41;	DB	2;	Length	275;
Best Local Similarity	77.8%;	Pred. No.	3.9;				
Matches	7;	Conservative	1;	Mismatches	1;	Indels	0;
Gaps	0;						

  

QY	2	FKFKFKFKF	10				
		:					
DB	210	FPYKFKFKF	218				

  

RESULT 2						
O86964	PRELIMINARY;	PRT;	208	AA.		
ID	O86964					
AC	O86964					
DT	01-NOV-1998 (TremBLrel. 08, Created)					
DT	01-NOV-1998 (TremBLrel. 08, Last sequence update)					
DT	01-NOV-1998 (TremBLrel. 08, Last annotation update)					
DE	DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).					
OS	Desulfovibrio sp. PT-2.					
OC	Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.					
RN	[1]					
RP	SEQUENCE FROM N.A.					

RC STRAIN=PT-2;  
 RX MEDLINE; 98269016.  
 RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
 RT "Phylogeny of dissimilatory sulfite reductases supports an early  
 RL origin of sulfate respiration."  
 RL J. Bacteriol. 180:2975-2982(1998).  
 DR EMBL; U58114; AAC24105.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 208 208  
 SQ SEQUENCE 208 AA; 23141 MW; 79E7999F CRC32;

Query Match 74.5%; Score 41; DB 2; Length 208;  
 Best Local Similarity 77.8%; Pred. No. 3.1;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | :|||||  
 Db 142 FPKFKFKF 150

RESULT 3  
 O86966 PRELIMINARY; PRT; 212 AA.  
 ID O86966;  
 AC O86966;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
 OS Desulfovibrio sp. P1B2.  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=P1B2.  
 RX MEDLINE; 98269016.  
 RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
 RT "Phylogeny of dissimilatory sulfite reductases supports an early  
 RL origin of sulfate respiration."  
 RL J. Bacteriol. 180:2975-2982(1998).  
 DR EMBL; U58116; AAC24107.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 212 212  
 SQ SEQUENCE 212 AA; 23441 MW; C90DA46D CRC32;

Query Match 74.5%; Score 41; DB 2; Length 212;  
 Best Local Similarity 77.8%; Pred. No. 3.1;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | :|||||  
 Db 142 FPKFKFKF 150

RESULT 4  
 O86970 PRELIMINARY; PRT; 223 AA.  
 ID O86970;  
 AC O86970;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
 OS Desulfohalobium sapovorans.  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33892;  
 RX MEDLINE; 98269016.  
 RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
 RT "Phylogeny of dissimilatory sulfite reductases supports an early  
 RL origin of sulfate respiration."  
 RL J. Bacteriol. 180:2975-2982(1998).  
 DR EMBL; U58120; AAC24109.1; -.

FT NON\_TER 1 1  
 FT NON\_TER 223 223  
 SQ SEQUENCE 223 AA; 24764 MW; FD37B9CC CRC32;

Query Match 74.5%; Score 41; DB 2; Length 223;  
 Best Local Similarity 77.8%; Pred. No. 3.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | :|||||  
 Db 142 FPKFKFKF 150

RESULT 5  
 O86974 PRELIMINARY; PRT; 367 AA.  
 ID O86974;  
 AC O86974;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
 OS Desulfohalobium latius.  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 43918;  
 RX MEDLINE; 98269016.  
 RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
 RT "Phylogeny of dissimilatory sulfite reductases supports an early  
 RL origin of sulfate respiration."  
 RL J. Bacteriol. 180:2975-2982(1998).  
 DR EMBL; U58124; AAC24097.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 367 367  
 SQ SEQUENCE 367 AA; 41543 MW; AACB5A60 CRC32;

Query Match 74.5%; Score 41; DB 2; Length 367;  
 Best Local Similarity 77.8%; Pred. No. 5;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | :|||||  
 Db 142 FPKFKFKF 150

RESULT 6  
 O86976 PRELIMINARY; PRT; 213 AA.  
 ID O86976;  
 AC O86976;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
 OS Desulfohalobium multivorans.  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33890;  
 RX MEDLINE; 98269016.  
 RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
 RT "Phylogeny of dissimilatory sulfite reductases supports an early  
 RL origin of sulfate respiration."  
 RL J. Bacteriol. 180:2975-2982(1998).  
 DR EMBL; U58126; AAC24101.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 213 213  
 SQ SEQUENCE 213 AA; 23637 MW; F50D1C25 CRC32;

Query Match 74.5%; Score 41; DB 2; Length 213;  
 Best Local Similarity 77.8%; Pred. No. 3.1;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFK 10

Db 142 FPKFKFKF 150

RESULT 7

ID O86978 PRELIMINARY; PRT; 221 AA.  
AC O86978;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DE DISIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
OS Desulfonema limicola.  
OC Bacteria; Proteobacteria; delta subdivision; Desulfonema.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 33961;  
RX MEDLINE; 98269016.  
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
RT "Phylogeny of dissimilatory sulfite reductases supports an early  
origin of sulfate respiration".  
RL J. Bacteriol. 180:2575-2582(1998).  
DR EMBL; U58128; AAC24099.1; -.  
FT NON\_TER 1  
FT NON\_TER 221  
SQ SEQUENCE 221 AA; 24396 MW; C81B1A78 CRC32;

Query Match 74.5%; Score 41; DB 2; Length 221;  
Best Local Similarity 77.8%; Pred. NO. 3.2;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFK 10

Db 142 FPKFKFKF 150

RESULT 8

ID Q92H18 PRELIMINARY; PRT; 471 AA.  
AC Q92H18;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DE DISIMILATORY SULFITE REDUCTASE SUBUNIT A.  
GN DSR4.  
OS Desulfotomaculum thermocisternum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Desulfotomaculum.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ST90;  
RA LARSEN O., LIEN T., BIRKELAND N.-K.;  
RT "Dissimilatory sulfite reductase from Archaeoglobus profundus and  
Desulfotomaculum thermocisternum: phylogenetic and structural  
implications from gene sequences".  
RL Extremophiles 0:0-0(1999).  
DR EMBL; AF043396; AAC96107.1; -.  
SQ SEQUENCE 471 AA; 54100 MW; 54C2B846 CRC32;

Query Match 74.5%; Score 41; DB 2; Length 471;  
Best Local Similarity 77.8%; Pred. NO. 6.2;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKFKFK 10

Db 244 FPKFKFKF 252

RESULT 9

RT "A proposed superfamily of transposase genes: transposon-like

O78874  
ID O78874 PRELIMINARY; PRT; 325 AA.  
AC O78874;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE NADH DEHYDROGENASE SUBUNIT II (FRAGMENT).  
GN ND2.  
OS Anolis equestris.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Lepidosauria;  
OC Squamata; Iguania; Iguanidae; Polychrotinae; Anolis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA JACKMAN T.R., LARSON A., DEQUEIROZ K., LOSOS J.B.;  
RT "Phylogenetic relationships and the tempo of early diversification in  
Anolis lizards".  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF055939; AAC27573.1; -.  
DR PFAM; PF00361; oxidored\_g1; 1.  
KW Mitochondrion.  
FT NON\_TER 1  
FT NON\_TER 325 AA; 36274 MW; A9FBEDDA CRC32;  
SQ SEQUENCE 325 AA; 36274 MW; A9FBEDDA CRC32;  
Query Match 74.5%; Score 41; DB 8; Length 325;  
Best Local Similarity 70.0%; Pred. NO. 4.5;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFKFKFKFK 10  
Db 292 KFKFKFKFK 301  
RESULT 10  
O50943 PRELIMINARY; PRT; 175 AA.  
ID O60943  
AC O60943;  
DT 01-AUG-1998 (TRENBLrel. 07, Created)  
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
DE 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE 20 KDA PROTEIN.  
OS Oxytricha fallax.  
OC Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;  
OC Oxytrichidae; Oxytricha.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBKARYONIDE 3.5;  
RX MEDLINE; 86079559.  
RA HERRICK G., CARTINHOOR S., DAWSON D., ANG D., SHEETS R., LEE A.,  
RA WILLIAMS K.;  
RT "Mobile elements bounded by C4A4 telomeric repeats in Oxytricha  
fallax".  
RL Cell 43:759-768(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBKARYONIDE 3.5;  
RX MEDLINE; 90185189.  
RA HUNTER D.J., WILLIAMS K., CARTINHOOR S., HERRICK G.;  
RT "Precise excision of telomere-bearing transposons during Oxytricha  
fallax macronuclear development".  
RL Genes Dev. 3:2101-2112(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBKARYONIDE 3.5;  
RA DOAK T.G., WILLIAMS K., HERRICK G.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBKARYONIDE 3.5;  
RX MEDLINE; 94134747.  
RA DOAK T.G., DOERDER F.P., JAHN C.L., HERRICK G.;  
RT "A proposed superfamily of transposase genes: transposon-like



AC O82294;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DE 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE PUTATIVE HIN1.  
 GN T32F12.16.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
 RT "Arabidopsis thaliana chromosome II BAC T32F12 genomic sequence.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC005314; AAC36175.1; -  
 SQ SEQUENCE 238 AA; 26825 MW; DE29EFF0 CRC32;

Query Match 67.3%; Score 37; DB 10; Length 238;  
 Best Local Similarity 60.0%; Pred. No. 15;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEFKFKFKF 10  
 | :|| :|||  
 Db 189 KLRFKLRPKF 198

RESULT 15  
 O33909  
 ID O33909 PRELIMINARY; PRT; 397 AA.  
 AC O33909;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last annotation update)  
 DE DISSIMILATORY SIROHEME-SULFITE REDUCTASE ALPHA SUBUNIT (EC 1.8.99.1).  
 GN DSRA.  
 OS Pyrobaculum islandicum.  
 OC Archaea; Crenarchaeota; Thermoproteales; Thermoproteaceae;  
 OC Pyrobaculum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 4184;  
 RA MOLITOR M., DAHL C., MOLITOR I., SCHAEFFER U., SPEICH N., HUBER R.,  
 RA DEUTZMANN R., TRUEPER H.G.;  
 RL Microbiology 143:0-0(1997).  
 DR EMBL: U75249; AAC46135.1; -  
 KW Oxidoreductase.  
 SQ SEQUENCE 397 AA; 44244 MW; 827A7545 CRC32;

Query Match 63.6%; Score 35; DB 1; Length 397;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FKFKFKFKF 10  
 | :||| ||  
 Db 184 FPKFKFKF 192

Search completed: September 7, 1999, 20:34:33  
 Job time: 19729 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:32 ; Search time 80.79 Seconds  
(without alignments)  
0.855 Million cell updates/sec

Title: US-09-124-280A-6  
Perfect score: 34  
Sequence: 1 KFLKKTLL 7  
Scoring table: BLOSUM62  
Searched: 106577 seqs, 9868381 residues  
Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	7	1	US-08-097-830E-6
2	34	100.0	7	2	US-08-456-112B-6
3	30	88.2	6	1	US-08-049-871-3
4	30	88.2	10	1	US-08-049-871-7
5	30	88.2	6	1	US-07-819-893-3
6	30	88.2	6	1	US-07-819-893-7
7	30	88.2	6	1	US-08-280-397-3
8	30	88.2	6	1	US-08-280-397-7
9	30	88.2	6	1	US-08-097-830E-30
10	30	88.2	10	2	US-08-218-026-48
11	30	88.2	10	2	US-08-653-632-48
12	30	88.2	6	2	US-08-456-112B-32
13	30	88.2	10	2	US-08-456-112B-36
14	30	88.2	10	3	PCT-US94-01234-46
15	28	82.4	341	1	US-08-423-564-5
16	28	82.4	890	1	US-08-145-006C-5
17	28	82.4	890	3	PCT-US94-00545-5
18	27	79.4	971	2	US-08-480-662-2
19	27	79.4	715	2	US-08-849-212-6
20	27	79.4	971	3	PCT-US96-03927-2
21	26	76.5	351	2	US-08-468-847B-16
22	25	73.5	7	1	US-08-049-871-1
23	25	73.5	10	1	US-08-049-871-2
24	25	73.5	11	1	US-08-049-871-6
25	25	73.5	271	1	US-07-914-282D-4
26	25	73.5	7	1	US-07-819-893-1
27	25	73.5	10	1	US-07-819-893-2
28	25	73.5	11	1	US-07-819-893-6
29	25	73.5	21	1	US-07-965-663A-1
30	25	73.5	21	1	US-07-965-663A-2
31	25	73.5	21	1	US-07-965-663A-3
32	25	73.5	22	1	US-07-965-663A-6
33	25	73.5	22	1	US-07-965-663A-7
34	25	73.5	23	1	US-07-965-663A-8
35	25	73.5	23	1	US-07-965-663A-9
36	25	73.5	23	1	US-07-965-663A-10
37	25	73.5	23	1	US-07-965-663A-11
38	25	73.5	24	1	US-07-965-663A-12
39	25	73.5	25	1	US-07-965-663A-13

Sequence 14, Appl  
Sequence 15, Appl  
Sequence 16, Appl  
Sequence 18, Appl  
Sequence 19, Appl  
Sequence 20, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-6  
; Sequence 6, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-6

Query Match 100.0%; Score 34; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKTLL 7  
Db 1 KFLKKTLL 7

RESULT 2  
US-08-456-112B-6  
; Sequence 6, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-6

Query Match 100.0%; Score 34; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKCTL 7  
Db 1 KFLKCTL 7

RESULT 3  
US-08-049-871-3  
Sequence 3, Application US/08049871  
Patent No. 5358933  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/049,871  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/658,744  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-049-871-3

Query Match 88.2%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 4  
US-08-049-871-7  
Sequence 7, Application US/08049871  
Patent No. 5358933  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/049,871  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/658,744  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-049-871-7

Query Match 88.2%; Score 30; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
Db 1 KFLKKT 6



```

Db      5 KFLKKT 10

RESULT 5
US-07-819-893-3
; Sequence 3, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: circular
; US-07-819-893-3

Query Match      88.2%; Score 30; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 KFLKKT 6
Db      1 KFLKKT 6

RESULT 6
US-07-819-893-7
; Sequence 7, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,397
; FILING DATE: 07/26/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/819,893
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-819-893-7

Query Match      88.2%; Score 30; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 KFLKKT 6
Db      5 KFLKKT 10

RESULT 7
US-08-280-397-3
; Sequence 3, Application US/08280397
; Patent No. 5589459
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,397
; FILING DATE: 07/26/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/819,893
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002a
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 ;  
 US-08-280-397-3

Query Match 88.2%; Score 30; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 Db 1 KFLKKT 6

RESULT 8  
 US-08-280-397-7  
 ; Sequence 7, Application US/08280397  
 ; Patent No. 5589459  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
 ; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
 ; TITLE OF INVENTION: Prevention and Treatment of Septic  
 ; TITLE OF INVENTION: Shock  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/280,397  
 ; FILING DATE: 07/25/94  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/819,893  
 ; FILING DATE: 01/16/92  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-002A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 10 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ;  
 US-08-280-397-7

Query Match 88.2%; Score 30; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 Db 5 KFLKKT 10

RESULT 9  
 US-08-097-830E-30  
 ; Sequence 30, Application US/08097830E  
 ; Patent No. 5652211  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: Peptides For Neutralizing The  
 ; TITLE OF INVENTION: Toxicity of Lipid A  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/097,830E  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 30:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 ;  
 US-08-097-830E-30

Query Match 88.2%; Score 30; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKT 6  
 Db 1 KFLKKT 6

RESULT 10  
 US-08-218-026-48  
 ; Sequence 48, Application US/08218026  
 ; Patent No. 5786324  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gray, Beulah  
 ; APPLICANT: Hasegan, Judith R.  
 ; APPLICANT: Mayo, Kevin  
 ; TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
 ; TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
 ; TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
 ; NUMBER OF SEQUENCES: 60  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merchant & Gould  
 ; STREET: 3100 No. 5786324west Center  
 ; CITY: Minneapolis  
 ; STATE: MN  
 ; COUNTRY: USA  
 ; ZIP: 55402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/218,026  
; FILING DATE: 24-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kowalchuk, Katherine M.  
; REGISTRATION NUMBER: 36,848  
; REFERENCE/DOCKET NUMBER: 600.286US01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-332-5300  
; TELEFAX: 612-332-9081  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-218-026-48

Query Match 88.2%; Score 30; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.2; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKKT 6  
Db 5 KFLKKT 10

RESULT 11  
US-08-653-632-48  
; Sequence 48, Application US/08653632  
; Patent No. 5830860  
; GENERAL INFORMATION:  
; APPLICANT: GRAY, Beulah  
; APPLICANT: HASEMAN, Judith R.  
; APPLICANT: MAYO, Kevin  
; TITLE OF INVENTION: PEPTIDES WITH BACTERICIDAL AND ENDOTOXIN NEUTRALIZING  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 5830860west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/653,632  
; FILING DATE: 24-MAR-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/218026  
; FILING DATE: 24-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kettelberger, Denise M  
; REGISTRATION NUMBER: 33,924  
; REFERENCE/DOCKET NUMBER: 600.286US11  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5268  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
; US-08-653-632-48

Query Match 88.2%; Score 30; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.2; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKKT 6  
Db 5 KFLKKT 10

RESULT 12  
US-08-456-112B-32  
; Sequence 32, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-32

Query Match 88.2%; Score 30; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKKT 6  
Db 1 KFLKKT 6

RESULT 13  
US-08-456-112B-36  
; Sequence 36, Application US/08456112B

; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Poito, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-36

Query Match 88.2%; Score 30; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
Db 5 KFLKKT 10

RESULT 14  
PCT-US94-01234-46  
; Sequence 46, Application PC/TUS9401234  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE  
; TITLE OF INVENTION: BINDING SITES  
; NUMBER OF SEQUENCES: 76  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01234  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/084,542  
; FILING DATE: 28-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/012,566  
; FILING DATE: 02-FEB-1993  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
PCT-US94-01234-46

Query Match 88.2%; Score 30; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
Db 5 KFLKKT 10

RESULT 15  
US-08-423-564-5  
; Sequence 5, Application US/08423564  
; Patent No. 5654417  
; Patent No. 5654417 5654417  
; GENERAL INFORMATION:  
; APPLICANT: Tarr, Phillip I.  
; APPLICANT: Bilge, Sima S.  
; APPLICANT: Vary, Jr., James C.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES FOR DETECTING E. coli O157:H7  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson, Kindness  
; STREET: 1420 5th Ave., Suite 2800  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/423,564  
; FILING DATE: 14-APR-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Broderick, Thomas F.  
; REGISTRATION NUMBER: 31,332  
; REFERENCE/DOCKET NUMBER: CHOR-1-8116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 682-8100  
; TELEFAX: (206) 224-0779  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 341 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: amino acids encoded by  
; DESCRIPTION: nucleotides 1-1026 of SEQ ID NO:1  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
US-08-423-564-5

Query Match 82.4%; Score 28; DB 1; Length 341;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 7  
Db 206 OFIKKTL 212

Search completed: September 7, 1999, 22:38:33  
Job time: 7917 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:19 ; Search time 116.8 Seconds  
(without alignments)  
2.401 Million cell updates/sec

Title: US-09-124-280A-6  
Perfect score: 34  
Sequence: 1 KFLKRTL 7

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	743	2 T00634	hypothetical prote
2	30	88.2	305	2 B71643	probable integrase
3	30	88.2	145	2 C49683	pts 3'-region prot
4	30	88.2	764	2 T01493	high affinity pota
5	29	85.3	248	2 S58093	probable multicata
6	29	85.3	4868	2 B54161	ryanodine-binding
7	29	85.3	322	2 A70189	hypothetical prote
8	29	85.3	662	2 S42826	probable ATPase -
9	29	85.3	141	2 C69339	potassium channel
10	29	85.3	507	2 T00871	probable 7-ethoxyc
11	29	85.3	180	2 JX0347	acrosomal major pr
12	28	82.4	307	1 WNV38	PE-38 protein - Or
13	28	82.4	165	2 S75527	H+-transporting AT
14	28	82.4	234	2 A64684	ribosomal protein
15	28	82.4	234	2 A71835	30s ribosomal prot
16	28	82.4	4872	2 S27272	ryanodine receptor
17	28	82.4	4859	2 S74173	ryanodine receptor
18	28	82.4	4869	2 S66572	ryanodine receptor
19	28	82.4	173	2 S44283	pail repressor hom
20	28	82.4	287	2 B64201	hypothetical prote
21	28	82.4	237	2 S73504	degv protein homol
22	28	82.4	299	2 H64354	hypothetical prote
23	28	82.4	232	2 A64366	hypothetical prote
24	28	82.4	1576	2 T03277	pol. protein - yeas
25	28	82.4	1355	2 S51995	probable ATPase (E
26	28	82.4	386	2 S51436	probable membrane
27	28	82.4	2024	2 A34103	centrosome autoant
28	28	82.4	890	2 A48753	NFAT transcripition
29	27	79.4	249	1 A70360	3',5'-cyclic-nucle
30	27	79.4	1192	1 ISKFAS	DNA topoisomerase
31	27	79.4	434	2 C71418	cytochrome P450 -
32	27	79.4	1094	2 S22573	DNA-directed DNA p
33	27	79.4	1191	2 S27329	DNA topoisomerase
34	27	79.4	410	2 C39088	contrapsin precurs
35	27	79.4	715	2 B41842	lysine decarboxyla
36	27	79.4	202	2 B71509	probable phosphata
37	27	79.4	659	2 S65976	hypothetical prote
38	27	79.4	493	2 PS0157	MDM10 protein - ye
39	27	79.4	1435	2 S69632	regulatory protein

#### ALIGNMENTS

##### RESULT 1

T00634

hypothetical protein H\_DJ0897G10.1 - human

C:Species: Homo sapiens (man)

C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 01-Feb-1999

C:Accession: T00634

R:Kalicki, J.; Elliott, G.

submitted to the EMBL Data Library, January 1998

A:Description: The sequence of Homo sapiens PAC clone DJ0897G10.

A:Reference number: Z14194

A:Accession: T00634

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-743 <KAL>

A:Cross-references: EMBL:AC004082; NID:g2822160; PID:g2822161

C:Genetics:

A:Map position: 7q11.23-q21.1

A:Introns: 36/2; 59/1; 130/3; 172/2; 203/3; 227/3; 288/1; 353/1; 414/1; 461/2; 486/1;

A:Note: H\_DJ0897G10.1

Query Match 94.1%; Score 32; DB 2; Length 743;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKRTL 7

Db 706 KFMKRTL 712

##### RESULT 2

B71643

probable integrase/recombinase ripx (xerc) RP817 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 21-Nov-1998

C:Accession: B71643

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630

A:Accession: B71643

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-305 <AND>

A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PID:el343087; PID:g386134

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: xerc; RP817

Query Match 88.2%; Score 30; DB 2; Length 305;  
Best Local Similarity 85.7%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKRTL 7

Db 92 KFLKRTL 98

##### RESULT 3

C49683

pts 3'-region protein 2 - Mycoplasma capricolium (SGC3) (fragment)

C:Species: Mycoplasma capricolum  
C:Date: 26-May-1995 #sequence\_revision 26-May-1995 #text\_change 09-Sep-1997  
C:Accession: C49683  
R:Zhu, P.P.; Reizer, J.; Reizer, A.; Peterkofsky, A.  
J. Biol. Chem. 268, 26531-26540, 1993  
A:Title: Unique monocistronic operon (ptsH) in Mycoplasma capricolum encoding the phosphorization of ptsH.  
A:Reference number: A49683; MUID:94075343  
A:Accession: C49683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-145 <ZHU>  
A:Cross-references: GB:L22432; NID:9435095; PID:9435099  
C:Genetics:  
A:Genetic code: SGC3

Query Match 88.2%; Score 30; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKKT 6  
Db 17 KFLKKT 22  
RESULT 4  
T01493  
high affinity potassium transporter protein homolog F107.17 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
C:Accession: T01493  
R:Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li, R.; D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N. submitted to the EMBL Data Library, June 1998  
A:Description: Arabidopsis thaliana chromosome 1 BAC F107 sequence.  
A:Reference number: Z14334  
A:Accession: T01493  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-764 <VYS>  
A:Cross-references: EMBL:AC003671; NID:92833627; PID:g3176686  
C:Genetics:  
A:Map position: 1  
A:Introns: 12/3; 178/1; 282/3; 409/1  
A:Note: F107.17

Query Match 88.2%; Score 30; DB 2; Length 764;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKKT 6  
Db 243 KFLKKT 248  
RESULT 5  
S58093  
probable multicatalytic endopeptidase complex (EC 3.4.99.46) chain SPA13C5.01c - fission  
N:Alternate names: protein SPA13C5.01c; protein SPA31A2.17c  
C:Species: Schizosaccharomyces pombe  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 29-May-1998  
C:Accession: S58093; S58109  
R:Devlin, K.; Churcher, C.M. submitted to the EMBL Data Library, July 1995  
A:Reference number: S58093  
A:Accession: S58093  
A:Molecule type: DNA  
A:Residues: 1-248 <DEV>  
A:Cross-references: EMBL:Z50112; NID:g908889; PID:g908890  
A:Accession: S58109  
A:Molecule type: DNA

A:Residues: 140-248 <DEW>  
A:Cross-references: EMBL:Z50113; NID:g914878; PID:g914895  
C:Superfamily: multicatalytic endopeptidase complex chain C9  
C:Keywords: hydrolase; proteinase

Query Match 85.3%; Score 29; DB 2; Length 248;  
Best Local Similarity 85.7%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKKT 7  
Db 194 KFLSKTL 200

RESULT 6  
B54161  
ryanodine-binding protein beta form - bullfrog  
C:Species: Rana catesbeiana (bullfrog)  
C:Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 29-Jan-1999  
C:Accession: B54161  
R:Oyamada, H.; Murayama, T.; Takagi, T.; Iino, M.; Iwabe, N.; Miyata, T.; Ogawa, Y.; J. Biol. Chem. 269, 17206-17214, 1994  
A:Title: Primary structure and distribution of ryanodine-binding protein isoforms of A:Reference number: A54161; MUID:94274714  
A:Accession: B54161  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-4868 <OVA>  
A:Cross-references: GB:D21071; NID:gl856973; PID:d1005176; PID:g538247  
C:Superfamily: ryanodine receptor

Query Match 85.3%; Score 29; DB 2; Length 4868;  
Best Local Similarity 85.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKKT 7  
Db 2822 KFLKKT 2828

RESULT 7  
A70189  
hypothetical protein BB0714 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998  
C:Accession: A70189  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh son, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: A70189  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-322 <KLE>  
A:Cross-references: GB:AE001171; GB:AE000783; NID:g2688640; PID:g2688650; TIGR:BB0714  
A:Experimental source: strain B31  
C:Superfamily: hypothetical protein BB0714; tetratricopeptide repeat homology  
F:122-155/Domain: tetratricopeptide repeat homology <Ttr1>  
F:161-193/Domain: tetratricopeptide repeat homology #status atypical <Ttr2>  
F:194-227/Domain: tetratricopeptide repeat homology <Ttr3>  
F:240-273/Domain: tetratricopeptide repeat homology <Ttr4>

Query Match 85.3%; Score 29; DB 2; Length 322;  
Best Local Similarity 85.7%; Pred. No. 43;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKKT 7



Db 143 KFLKKAL 149  
|||||

## RESULT 8

S42826  
probable ATPase - fluke (Schistosoma mansoni)

C:Species: Schistosoma mansoni  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 04-Sep-1998  
C:Accession: S42826

R:Kunz, W.  
submitted to the EMBL Data Library, February 1994

A:Reference number: S42826

A:Accession: S42826

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-662 <KUN>

A:Cross-references: EMBL:229947; NID:q453977; PID:q453978

C:Superfamily: cell division protein ftsH; FtsH/SEC18/CDC48-type ATP-binding domain homology

C:Keywords: ATP; P-loop

F:180-387/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VATP>

F:206-213/Region: nucleotide-binding motif A (P-loop)

## Query Match

Best Local Similarity 85.3%; Score 29; DB 2; Length 662;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKKTLL 7

Db 137 FLKKTLL 142

|||||

## RESULT 9

C69339

potassium channel homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998

C:Accession: C69339

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343

A:Accession: C69339

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-141 <KLE>

A:Cross-references: GB:AE001055; GB:AE000782; NID:g2689378; PID:g2649899; TIGR:AF0715

## Query Match

Best Local Similarity 85.3%; Score 29; DB 2; Length 141;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKKTLL 7

Db 134 FLKKTLL 139

|||||

## RESULT 10

T00871

probable 7-ethoxycoumarin O-deethylase F17K2.11 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999

C:Accession: T00871

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, March 1998

A:Description: Arabidopsis thaliana chromosome II BAC F17K2 genomic sequence.

A:Reference number: 214207

A:Accession: T00871

A:Status: preliminary; translated from GB/EMBL/DDBB

A:Molecule type: DNA

A:Residues: 1-507 <ROU>

A:Cross-references: EMBL:AC003680; NID:g2979540; PID:g2979550

C:Genetics:

A:Map position: II

A:Introns: 123/2; 300/3; 393/3

A:Note: F17K2.11

## Query Match

Best Local Similarity 85.3%; Score 29; DB 2; Length 507;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKTL 7

Db 123 KFLKKTI 129

|||||

## RESULT 11

JX0347

acrosomal major protein M7 - blue mussel

C:Species: Mytilus edulis (blue mussel)

C:Date: 22-Apr-1995 #sequence\_revision 26-May-1995 #text\_change 26-May-1995

C:Accession: JX0347

R:Takagi, T.; Nakamura, A.; Deguchi, R.; Kyoizuka, K.

J. Biochem. 116, 598-605, 1994

A:Title: Isolation, characterization, and primary structure of three major proteins o

A:Reference number: JX0349

A:Accession: JX0347

A:Molecule type: protein

A:Residues: 1-180 <TAK>

A:Experimental source: sperm

C:Comment: This protein is one of the three major proteins and shows egg vitelline co

F:54-65,84-176,153-168/disulfide bonds: #status predicted

## Query Match

Best Local Similarity 85.3%; Score 29; DB 2; Length 180;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKKTLL 7

Db 102 FLKKTLL 107

|||||

## RESULT 12

WMNV38

PE-38 protein - Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus

C:Species: Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OpNPV

C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 13-Sep-1997

C:Accession: A42191

R:Theilmann, D.A.; Stewart, S.

Virology 187, 97-106, 1992

A:Title: Tandemly repeated sequence at the 3' end of the IE-2 gene of the baculovirus

A:Reference number: A42191; MUID:92142537

A:Accession: A42191

A:Molecule type: DNA

A:Residues: 1-307 <THE>

A:Cross-references: GB:M83827; NID:g332540; PID:g332542

C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein

C:Keywords: DNA binding; early protein

## Query Match

Best Local Similarity 82.4%; Score 28; DB 1; Length 307;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKKTL 6

Db 110 KFLKKTI 115

|||||

## RESULT 13

S75527  
H+-transporting ATP synthase (EC 3.6.1.34) epsilon chain - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: ATP synthase e chain; protein slr1330  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
R:Accession: S75527; S17754; S14859  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S75527  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-165 <KAN>  
A:Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PID:d1018821; PID:g1653172  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996  
R:Lill, H.; Nelson, N.  
Plant Mol. Biol. 17, 641-652, 1991  
A:Title: The atp1 and atp2 operons of the cyanobacterium *Synechocystis* sp. PCC 6803.  
A:Reference number: S17745; MUID:92003679  
A:Accession: S17754  
A:Molecule type: DNA  
A:Residues: 30-165 <LIL>  
A:Cross-references: EMBL:X58129; NID:g47515; PID:g47517  
C:Genetics:  
A:Gene: atpE  
C:Superfamily: H+-transporting ATP synthase epsilon chain  
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex

Query Match 82.4%; Score 28; DB 2; Length 165;  
Best Local Similarity 85.7%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
|||||  
Db 26 KFLKMTLL 32

## RESULT 14

A64684  
ribosomal protein S3 - *Helicobacter pylori* (strain 26695)  
C:Species: *Helicobacter pylori*  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 31-Oct-1997  
C:Accession: A64684  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A:Reference number: A64520; MUID:97394467  
A:Accession: A64684  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-234 <TOM>  
A:Cross-references: GB:AE000633; GB:AE000511; NID:g2314452; PID:g2314475; TIGR:HP1313  
C:Superfamily: *Escherichia coli* ribosomal protein S3

Query Match 82.4%; Score 28; DB 2; Length 234;  
Best Local Similarity 85.7%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
|||||  
Db 41 KFLKKEEL 47

## RESULT 15

A71835  
30s ribosomal protein S3 - *Helicobacter pylori* (strain J99)  
C:Species: *Helicobacter pylori*  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 05-Mar-1999  
C:Accession: A71835  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*.  
A:Reference number: A71800; MUID:99120557  
A:Accession: A71835  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-234 <ARN>  
A:Cross-references: GB:AE001547; GB:AE001439; NID:g4155811; PID:g4155825  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: rpsC  
C:Superfamily: *Escherichia coli* ribosomal protein S3

Query Match 82.4%; Score 28; DB 2; Length 234;  
Best Local Similarity 85.7%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
|||||  
Db 41 KFLKKEEL 47

Search completed: September 7, 1999, 23:06:21  
Job time: 2469 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:06 ; Search time 71.87 Seconds  
(without alignments)  
2.753 Million cell updates/sec

Title: US-09-124-280A-6

Perfect score: 34

Sequence: 1 KFLKKT 7

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	88.2	145	1 YPH2_MYCCA	P45614 mycoplasma
2	29	85.3	507	1 C7C3_ARATH	O64638 arabidopsis
3	29	85.3	248	1 PRC9_SCHPO	Q09682 schizosacch
4	29	85.3	662	1 YME1_SCHMA	P46508 schizosoma
5	28	82.4	1355	1 ATC3_YEAST	P39524 saccharomyc
6	28	82.4	521	1 ICEA_HUMAN	Q92851 homo sapien
7	28	82.4	233	1 NCAP_BUNGE	P69993 bunyavirus
8	28	82.4	307	1 PE38_NPVOP	P32512 oryza pseu
9	28	82.4	386	1 PEXD_YEAST	P80667 saccharomyc
10	28	82.4	234	1 R33_HELPY	P56010 helicobacte
11	28	82.4	287	1 Y011_MYCGE	P47257 mycoplasma
12	28	82.4	299	1 Y440_METJA	Q57882 methanococc
13	28	82.4	237	1 Y450_MYCPN	P75127 mycoplasma
14	28	82.4	232	1 Y529_METJA	Q57949 methanococc
15	28	82.4	173	1 YPIP_LACDL	P46543 lactobacill
16	27	79.4	971	1 CAS_HUMAN	P35060 homo sapien
17	27	79.4	412	1 CCA_SULSH	P77978 sulfolobus
18	27	79.4	410	1 COTR_CAVPO	P23233 cavia porce
19	27	79.4	715	1 DCLY_ECOLI	P23892 escherichia
20	27	79.4	1094	1 DPOD_PLAFK	P30315 plasmodium
21	27	79.4	2199	1 DPOE_SCHPO	P87154 schizosacch
22	27	79.4	493	1 MD10_YEAST	P18409 saccharomyc
23	27	79.4	346	1 RDS_YENLA	O42583 xenopus lae
24	27	79.4	1395	1 SP41_YEAST	P38904 saccharomyc
25	27	79.4	1192	1 TOP2_ASFB7	Q00942 african swi
26	27	79.4	1191	1 TOP2_ASFM2	P34203 african swi
27	27	79.4	659	1 YVBT_BACSU	P37484 bacillus su
28	26	76.5	1478	1 BCK1_YEAST	Q01389 saccharomyc
29	26	76.5	840	1 CC16_YEAST	P05798 saccharomyc
30	26	76.5	296	1 DPSD_CLOPA	Q46192 clostridium
31	26	76.5	650	1 GYRB_MYCPN	P23447 mycoplasma
32	26	76.5	1321	1 IF31_DICDI	O15818 dictyosteli
33	26	76.5	277	1 MCRA_ECOLI	P24200 escherichia
34	26	76.5	351	1 NOV_CHICK	P28686 gallus gall
35	26	76.5	353	1 NOV_COTJA	P23642 coturnix co
36	26	76.5	148	1 R19G_ASCSU	P24494 ascaris suu
37	26	76.5	364	1 RDS2_CHICK	O42282 gallus gall
38	26	76.5	201	1 TRI_PONLE	P05547 pontastacus
39	26	76.5	1823	1 VIT_ICHUN	Q91062 ichtthyomizo
40	26	76.5	299	1 Y060_MYCPN	P75042 mycoplasma
41	26	76.5	950	1 Y136_HUMAN	Q14149 homo sapien
42	26	76.5	246	1 Y419_ARCFU	O29828 archaeoglob
43	26	76.5	300	1 Y921_METJA	O58331 methanococc

ALIGNMENTS

```
RESULT 1
YPH2_MYCCA
ID YPH2_MYCCA STANDARD; PRT; 145 AA.
AC P45614;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN IN PTSH 3'REGION (ORF2) (FRAGMENT).
OS MYCOPLASMA CAPRICOLUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC CAPRICOLUM GROUP.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343 / KID;
RX MEDLINE: 94075343.
RA ZHU P.-P., REIZER J., REIZER A., PETERKOVSKY A.;
RT "unique monocistronic operon (ptsh) in Mycoplasma capricolum encoding
RT the phosphocarrier protein, Hpr, of the phosphoenolpyruvate:sugar
RT phosphotransferase system. Cloning, sequencing, and characterization
RT of ptsh.";
RL J. BIOL. CHEM. 268:26531-26540(1993).
CC - SIMILARITY: BELONGS TO THE RLJ FAMILY OF PSEUDOURIDINE SYNTHASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L22432; G435099;
DR PROSITE; PS01129; PSI_RLU; PARTIAL.
DR KW HYPOTHETICAL PROTEIN.
FT NON_TER 145 145
SQ SEQUENCE 145 AA; 17314 MW; 9291B799 CRC32;

Query Match 88.2%; Score 30; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 7.5; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 1 KFLKKT 6
Db 17 KFLKKT 22

RESULT 2
C7C3_ARATH
ID C7C3_ARATH STANDARD; PRT; 507 AA.
AC O64638;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CYTOCHROME P450 76C3 (EC 1.14.-.-).
GN CYP76C3 OR F17K2.11.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; EMERYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA ROUNSLEY S.D., KAUL S., LIN X., KETCHUM K.A., CROSBY M.L.,
RA BRANSON R.C., SYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
RA SOMERVILLE C.R., VENTER J.C.;
```

44 26 76.5 366 1 YA32\_METJA Q58438 methanococc  
45 26 76.5 320 1 YDJE\_BACSU O34768 bacillus su

RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AC003680; G2979550; -.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW OXIDOREDUCTASE; MONOOXYGENASE; HEME; MULTIGENE FAMILY.  
FT BINDING 443 443 HEME (BY SIMILARITY).  
SQ SEQUENCE 507 AA; 57285 MW; EEB44B27 CRC32;

Query Match 85.3%; Score 29; DB 1; Length 507;  
Best Local Similarity 71.4%; Pred. No. 43;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
DB 123 RFLKKTLL 129

RESULT 3  
PRC9\_SCHPO STANDARD; PRT; 248 AA.  
AC Q09682;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PUTATIVE PROTEASOME COMPONENT C9/Y13 (EC 3.4.99.46) (MACROPAIN  
DE SUBUNIT) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT).  
GN SPAC13C5.O1C OR SPAC31A2.17C.  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;  
OC SCHIZOSACCHAROMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA DEVLIN K.; CHURCHER C.M.; BARRELL B.G.; RAJANDREAN M.A.; WALSH S.V.;  
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH  
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT  
CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT  
CC PROTEOLYTIC ACTIVITY (POTENTIAL).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A. ALSO KNOWN AS THE  
CC PROTEASOME A-TYPE FAMILY. BELONGS TO THE C9 SUBFAMILY.  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Z50112; G908890; -.  
EMBL; Z50113; G914895; -.  
DR PROSITE; PS00388; PROTEASOME\_A; 1.  
DR PFAM; PF00227; proteasome; 1.  
DR HSP; P25156; 1PMA.  
KW PROTEASOME; HYDROLASE; PROTEASE.  
SQ SEQUENCE 248 AA; 27925 MW; EFBA8911 CRC32;

Query Match 85.3%; Score 29; DB 1; Length 248;

Best Local Similarity 85.7%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
DB 194 KFLSKTLL 200

RESULT 4  
YMEL\_SCHMA STANDARD; PRT; 662 AA.  
ID YMEL\_SCHMA  
AC P46508;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE YMEL PROTEIN HOMOLOG (EC 3.4.24.-).  
OS SCHISTOSOMA MANSONI (BLOOD FLUKE).  
OC EUKARYOTA; METAZOA; PLATYHELMINTHES; TREMATODA; DIGENEA; STRIGEIDIDA;  
OC SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LIBERIAN;  
RA MENRATH M.;  
RL THESIS (1994), HEINRICH-HEINE UNIVERSITY / DUESSELDORF, GERMANY.  
CC -!- FUNCTION: PUTATIVE ATP-DEPENDENT PROTEASE.  
CC -!- COFACTOR: BINDS AND REQUIRES A ZINC ATOM (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41 (ZINC  
CC METALLOPROTEASE).  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Z29947; G453978; -.  
DR PROSITE; PS00674; AAA; 1.  
DR PFAM; PF00004; AAA; 1.  
KW ATP-BINDING; HYDROLASE; METALLOPROTEASE; ZINC.  
FT NP\_BIND 206 213 ATP (POTENTIAL).  
FT METAL 425 425 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 426 426 BY SIMILARITY.  
FT METAL 429 429 ZINC (CATALYTIC) (BY SIMILARITY).  
SQ SEQUENCE 662 AA; 72939 MW; E8F6CD5D CRC32;

Query Match 85.3%; Score 29; DB 1; Length 662;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKKTLL 7  
DB 137 FLKKTLL 142

RESULT 5  
ATC3\_YEAST STANDARD; PRT; 1355 AA.  
ID ATC3\_YEAST  
AC P39524;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PROBABLE CALCIUM-TRANSPORTING ATPASE 3 (EC 3.6.1.38) (ENDOPLASMIC  
DE RETICULUM CA2+-ATPASE).  
GN DRS2 OR YAL026C OR FUN38.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE: 94067151.  
RA RIPAFTER T.L., VAUGHN G.P., WOOLFORD J.L. JR.:  
RT "DRS1 TO DRS7, novel genes required for ribosome assembly and  
RL function in Saccharomyces cerevisiae.";  
RM MOL. CELL. BIOL. 13:7901-7912(1993).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN-S288C / AB972;  
RX MEDLINE: 95249563.  
RA BUSSEY H., KABACK D.B., ZHONG W., VO D.T., CLARK M.W., FORTIN N.,  
RA HALL J., OUELLETTE B.F.F., KENG T., BARTON A.B., SU Y., DAVIES C.K.,  
RA STORMS R.K.;  
RT "The nucleotide sequence of chromosome I from Saccharomyces  
RT cerevisiae.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 92:3809-3813(1995).  
CC -|- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE  
CC HYDROLYSIS OF ATP COUPLED WITH THE TRANSPORT OF CALCIUM  
CC (BY SIMILARITY). SEEMS TO BE INVOLVED IN RIBOSOME ASSEMBLY.  
CC -|- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + ORTHOPHOSPHATE.  
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
CC RETICULUM.  
CC -|- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
CC (E1-E2 ATPASES).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L01795; G171114; -;  
DR EMBL: U12980; G595560; -;  
DR PIR: S30768; S30768.  
DR SGD: L0000526; DRS2.  
DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
DR PFAM: PF00122; E1-E2\_ATPase; 2.  
KW HYDROLASE; CALCIUM TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;  
FT MAGNESIUM; ATP-BINDING; ENDOPLASMIC RETICULUM.  
FT MOD\_RES 560 560 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 1355 AA; 15384 MW; 78D00FF8 CRC32;  
  
Query Match 82.4%; Score 28; DB 1; Length 1355;  
Best Local Similarity 85.7%; Pred. NO. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Oy 1 KFLKKTL 7  
Db 991 KFLKKLL 997  
  
RESULT 6  
ID ICEA\_HUMAN STANDARD: PRT; 521 AA.  
AC Q92851; Q99845;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE CASPASE-10 PRECURSOR (EC 3.4.22.-) (ICE-LINE APOPTOTIC PROTEASE 4)  
DE (APOPTOTIC PROTEASE MCH-4) (FAS-ASSOCIATED DEATH DOMAIN PROTEIN  
DE INTERLEUKIN-1B-CONVERTING ENZYME 2) (FLICE2).  
GN CASP10 OR MCH4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A. (SHORT FORM).  
RC TISSUE=T-CELL;  
RX MEDLINE: 96353838.  
RA FERNANDES-ALNEMRI T., ARMSTRONG R.C., KREBS J., SRINIVASULA S.M.,  
RA WANG L., BULLRICH F., FRITZ L.C., TRAPANI J.A., TOMASELLI R.J.,

RA LITWACK G., ALNEMRI E.S.;  
RT "In vitro activation of CPP32 and Mch3 by Mch4, a novel human  
RT apoptotic cysteine protease containing two FADD-like domains.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 93:7464-7469(1996).  
[2]  
RN SEQUENCE FROM N.A. (LONG FORM).  
RP MEDLINE: 97197836.  
RX VINCENZ C., DIXIT V.M.;  
RA "Fas-associated death domain protein interleukin-1beta-converting  
RT enzyme 2 (FICE2), an ICE/Ced-3 homologue, is proximally involved in  
RT CD95- and p55-mediated death signalling.";  
RL J. BIOL. CHEM. 272:6578-6583(1997).  
[3]  
RN PARTIAL SEQUENCE, AND PROCESSING.  
RP MEDLINE: 97121412.  
RX SRINIVASULA S.M., AHMAD M., FERNANDES-ALNEMRI T., LITWACK G.,  
RA ALNEMRI E.S.;  
RA "Molecular ordering of the Fas-apoptotic pathway: the Fas/APO-1  
RT protease Mch5 is a CrmA-inhibitable protease that activates multiple  
RT Ced-3/ICE-like cysteine proteases.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 93:14486-14491(1996).  
CC -|- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES  
CC RESPONSIBLE FOR APOPTOSIS EXECUTION. RECRUITED TO BOTH FAS- AND  
CC TNFR-1 RECEPTORS IN A FADD DEPENDENT MANNER. MAY PARTICIPATE IN  
CC THE GRANZYME B APOPTOTIC PATHWAYS. CLEAVES AND ACTIVATES CASPASE-  
CC 3, -4, -6, -7, -8, AND -9. HYDROLYZES THE SMALL- MOLECULE  
CC SUBSTRATES, TYR-VAL-ALA-ASP-I-AMC AND ASP-GLU-VAL-ASP-I-AMC.  
CC -|- SUBUNIT: HETERODIMER OF A 23/17 KD (P23/17) DEPENDING ON THE  
CC SPLICING EVENTS AND A 12 KD (P12) SUBUNIT.  
CC -|- TISSUE SPECIFICITY: DETECTABLE IN MOST TISSUES. LOWEST EXPRESSION  
CC IS SEEN IN BRAIN, KIDNEY, PROSTATE, TESTIS, AND COLON.  
CC -|- DOMAIN: CONTAINS TWO DEATH EFFECTOR DOMAINS (ALSO CALLED MORT  
CC MODULES), WHICH ARE INVOLVED IN BINDING TO THE ADAPTOR MOLECULE  
CC FADD.  
CC -|- PTM: CLEAVAGE BY GRANZYME B AND AUTOCATALYTIC ACTIVITY GENERATE  
CC THE TWO ACTIVE SUBUNITS.  
CC -|- ALTERNATIVE PRODUCTS: A LONG FORM (SHOWN HERE) AND A SHORT FORM  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE  
CC CASPASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U60519; G1498324; -;  
DR EMBL: U86214; G1835779; -;  
DR MIN: 601762; -;  
DR PROSITE: PS01121; CASPASE\_HIS; 1.  
DR PROSITE: PS01122; CASPASE\_CYS; 1.  
DR PFAM: PF00653; ICE\_P10; 1.  
DR PFAM: PF00656; ICE\_P20; 1.  
DR PFAM: PF01335; DED; 2.  
DR HSSP: P42574; 1PAU.  
KW HYDROLASE; THIOL PROTEASE; APOPTOSIS; ZYMOGEN; ALTERNATIVE SPLICING.  
FT PROPEP 1 219  
FT CHAIN 220 415 CASPASE-10 SUBUNIT P23/17.  
FT CHAIN 416 521 CASPASE-10 SUBUNIT P12.  
FT CHAIN 20 96 DEATH EFFECTOR DOMAIN.  
FT DOMAIN 115 187 DEATH EFFECTOR DOMAIN.  
FT ACT\_SITE 358 358 BY SIMILARITY.  
FT ACT\_SITE 401 401 BY SIMILARITY.  
FT VARSPLIC 229 271 MISSING (IN SHORT FORM).  
FT VARSPLIC 473 521 MLKLEKTEIMETRGKRTVWGAQISATSLPTAISAQTPRP  
PMRWSSVS -> HEDILSILTAVNDVSRVQGTGKQ  
MPQAFTRLRKLVFVPLDLSI (IN SHORT FORM).  
FT CONFLICT 68 68 G -> E (IN REF. 1).  
SQ SEQUENCE 521 AA; 58878 MW; D4D64D28 CRC32;

Query Match 82.4%; Score 28; DB 1; Length 521;  
Best Local Similarity 71.4%; Pred. No. 74;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
|||:|:  
DB 475 KFEKTM 481

RESULT 7  
NCAP\_BUNGE STANDARD; PRT; 233 AA.  
AC P16993;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-AUG-1990 (REL. 15, LAST ANNOTATION UPDATE)  
DE NUCLEOCAPSID PROTEIN (NUCLEOPROTEIN).  
GN N.  
OS BUNYAVIRUS GERMISTON.  
OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; BUNYAVIRIDAE; BUNYAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88019678.  
RA GERBAUD S., VIALAT P., PARDIGON N., WYCHOWSKI C., GIRARD M.,  
RA BOULOY M.;  
RT "The S segment of the GERMISTON virus RNA genome can code for three  
RT proteins.";  
RL VIRUS RES. 8:1-13(1987).  
CC -1- SUBCELLULAR LOCATION: INTERNAL PROTEIN OF VIRUS PARTICLE.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M19420; G210758; -  
DR PFAM; PF00952; Buoya\_nucleocap; 1.  
KW NUCLEOCAPSID; RNA-BINDING.  
SQ SEQUENCE 233 AA; 26703 MW; B4393642 CRC32;

Query Match 82.4%; Score 28; DB 1; Length 233;  
Best Local Similarity 71.4%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKTLL 7  
|||:|:  
DB 175 KFEKTM 181

RESULT 8  
PE38\_NPVOP STANDARD; PRT; 307 AA.  
AC P22512;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MAJOR IMMEDIATE EARLY PROTEIN (PE-38).  
GN PE38.  
OS ORGYIA PSEUDOTSUGATA MULTICAPSID POLYHEDROSIS VIRUS (OPNPV).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;  
OC NUCLEOPOLYHEDROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92142537.  
RA THEILMANN D.A., STEWART S.;  
RT "Randomly repeated sequence at the 3' end of the IE-2 gene of the  
RT baculovirus Orgyia pseudotsugata multicaapsid nuclear polyhedrosis  
RT virus is an enhancer element.";

RL VIROLOGY 187:97-106(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97271300.  
RA AHRENS C.H., RUSSELL R.R., FUNK C.J., EVANS J., HARWOOD S.,  
RA ROHRMANN G.F.;  
RT "The sequence of the Orgyia pseudotsugata multicaapsid nuclear  
RT polyhedrosis virus genome.";  
RL VIROLOGY 229:381-398(1997).  
CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M83827; G332542; -  
DR EMBL; U75930; G1911398; -  
DR PIR; A42191; WMNV38.  
DR PROSITE; P500518; ZINC\_FINGER\_C3HC4; 1.  
DR PFAM; PF00097; ZF-C3HC4; 1.  
KW EARLY PROTEIN; ZINC-FINGER; DNA-BINDING.  
FT ZN.FING 39 91 C3HC4-TYPE.  
SQ SEQUENCE 307 AA; 34714 MW; E712CE99 CRC32;

Query Match 82.4%; Score 28; DB 1; Length 307;  
Best Local Similarity 83.3%; Pred. No. 43;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKKT 6  
||:|:  
DB 110 KFIKKT 115

RESULT 9  
PEXD\_YEAST STANDARD; PRT; 386 AA.  
AC P80667;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE PEROXISOMAL MEMBRANE PROTEIN PAS20 (PEROXIN-13).  
GN PEX13 OR PAS20 OR YLR191W OR L9470.1.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97011156.  
RA ELGERSMA Y., KWAST L., KLEIN A., VOORN-BROUWER T., VAN DEN BERG M.,  
RA TABAK H.F., DISTEL B.;  
RT "The SH3 domain of the Saccharomyces cerevisiae peroxisomal membrane  
RT protein Pex13p functions as a docking site for Pex5p, a mobile  
RT receptor for the import PTS1-containing proteins.";  
RL J. CELL BIOL. 135:97-109(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
RA WILSON R., WATERSTON R.;  
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN-S288C;

```

RX MEDLINE: 97011157.
RA ERDMANN R., BLOBEL G.;
RT "Identification of Pex13p a peroxisomal membrane receptor for the
RL PTS1 recognition factor.";
RL J. CELL BIOL. 135:111-121(1996).
CC -!- FUNCTION: COMPONENT OF THE PEROXISOMAL TRANSLLOCATION MACHINERY
CC WITH PEX14 AND PEX17. INTERACTS WITH THE PTS1 RECEPTOR
CC (PAS10/PEX5).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: S82971; G1836116; -
DR EMBL: U37420; G1019899; -
DR EMBL: U17246; G577193; -
DR EMBL: U14913; G2340981; -
DR SGD: L0003024; PAS20.
DR PROSITE: PS50002; SH3; 1.
DR PFAM: PF00018; SH3; 1.
DR HSP: Q06187; 1AWN.
KW PEROXISOME; TRANSPORT; PROTEIN TRANSPORT; TRANSMEMBRANE; SH3 DOMAIN.
FT DOMAIN 1 263
FT TRANSMEM 264 280
FT DOMAIN 281 386
FT DOMAIN 306 372
FT SH3.
SQ SEQUENCE 386 AA; 42706 MW; 4265AD81 CRC32;
Query Match 82.4%; Score 28; DB 1; Length 386;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KFLKRTL 7
DB 205 KFLKRTL 211
||| | |
RESULT 10
RS3_HELPY STANDARD; PRT; 234 AA.
AC P56010;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE 30S RIBOSOMAL PROTEIN S3.
GN RPS3 OR HP1313
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;
OC HELICOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-26695 / ATCC 700392;
RX MEDLINE: 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL NATURE 388:539-547(1997).
CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN THE BINDING OF INITIATOR

```

```

CC MET-TRNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000633; G2314475; -
DR TIGR: HP1313; -
DR PROSITE: PS00548; RIBOSOMAL_S3; 1.
DR PFAM: PF00113; KH-domain; 1.
DR PFAM: PF00189; S3_C; 1.
DR PFAM: PF00417; S3_N; 1.
KW RIBOSOMAL PROTEIN; TRNA-BINDING.
SQ SEQUENCE 234 AA; 26380 MW; A3D85FDB CRC32;
Query Match 82.4%; Score 28; DB 1; Length 234;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KFLKRTL 7
DB 41 KFLKREL 47
||| | |
RESULT 11
Y011_MYCGE STANDARD; PRT; 287 AA.
AC P47257; Q49362;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MG011.
GN MG011.
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE: 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITZCHMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 158-255 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE: 94075230.
RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. BACTERIOL. 175:7918-7930(1993).
CC -!- SIMILARITY: BELONGS TO THE ATP-GRASP SUPERFAMILY; RIMK FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U39679; G1045679; -

```

DR EMBL: U02257; G405920; -  
 DR TIGR: MG011; -  
 KW HYPOTHETICAL PROTEIN.  
 FT CONFLICT 158 T -> P (IN REF. 2).  
 FT CONFLICT 255 A -> R (IN REF. 2).  
 SQ SEQUENCE 287 AA; 33433 MW; 9E2DA74F CRC32;

Query Match 82.4%; Score 28; DB 1; Length 287;  
 Best Local Similarity 85.7%; Pred. No. 41;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTL 7  
 |||||

DB 279 KFLKKKL 285

RESULT 12  
 Y440\_METJA STANDARD; PRT; 299 AA.  
 AC Q57882;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MJ0440.

OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43087;  
 RX MEDLINE; 96337999.

RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA KUTLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL SCIENCE 273:1058-1073(1996).

CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -I- SIMILARITY: IN THE N-TERMINAL, TO P.WOESLI HYPOTHETICAL 12.3 KD  
 CC PROTEIN IN FUS 3'REGION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL: U67495; G1591144; -

DR TIGR: MJ0440; -

KW HYPOTHETICAL PROTEIN: TRANSMEMBRANE.

FT TRANSMEM 54 74 POTENTIAL.

FT TRANSMEM 89 109 POTENTIAL.

FT TRANSMEM 122 142 POTENTIAL.

FT TRANSMEM 149 169 POTENTIAL.

FT TRANSMEM 172 192 POTENTIAL.

FT TRANSMEM 208 228 POTENTIAL.

FT TRANSMEM 243 263 POTENTIAL.

SQ SEQUENCE 299 AA; 34301 MW; 32E37441 CRC32;

Query Match 82.4%; Score 28; DB 1; Length 299;  
 Best Local Similarity 85.7%; Pred. No. 42;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTL 7  
 |||||

DB 190 KFLDKTL 196

RESULT 13

Y450\_MYCPN STANDARD; PRT; 237 AA.

AC P75127;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN MG450 HOMOLOG.

OS MYCOPLASMA PNEUMONIAE.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;

OC MYCOPLASMATACEAE; MYCOPLASMA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 29342 / M129;

RX MEDLINE; 97105885.

RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,

RA HERMANN R.

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 pneumoniae.";  
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL: AE000019; G1673841; -

DR HYPOTHETICAL PROTEIN.

KW SEQUENCE 237 AA; 26839 MW; 407D0E1D CRC32;

Query Match 82.4%; Score 28; DB 1; Length 237;

Best Local Similarity 85.7%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKKTL 7

|||||

DB 167 KFLKKQL 173

RESULT 14

Y529\_METJA STANDARD; PRT; 232 AA.

AC Q57949;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN MJ0529.

GN MJ0529.

OS METHANOCOCCUS JANNASCHII.

OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;

OC METHANOCOCCUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;

RX MEDLINE; 96337999.

RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,

RA KUTLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,

RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,

RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,

RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,

RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,

RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus



Search completed: September 7, 1999, 23:50:07  
Job time: 1946 sec

```
RT jannaschii."
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: WEAK, TO M.JANNASCHII MJ0173.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67502; G1591232; -
DR TIGR: M30329; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 232 AA; 27091 MW; 7E4228B8 CRC32;

Query Match      82.4%; Score 28; DB 1; Length 232;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 6
DB 197 KFMKKT 202
   ||:||||

RESULT 15
YPIP_LACDCL
ID YPIP_LACDCL STANDARD; PRT; 173 AA.
AC P46543;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 19.8 KD PROTEIN IN PEPI 3'REGION.
OS LACTOBACILLUS DELBRUECKII (SUBSP. LACTIS).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;
OC LACTOBACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 7290;
RX MEDLINE: 94297735,
RA KLEIN J.R., SCHMIDT U., PLAPP R.;
RT "Cloning, heterologous expression, and sequencing of a novel proline
RT iminopeptidase gene, pepI, from Lactobacillus delbrueckii subsp.
RT lactis DSM 7290."
RL MICROBIOLOGY 140:1133-1139(1994).
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS PA1A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z26948; G482922; -
DR PFAM: PF00583; Acetyltransf; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 173 AA; 19819 MW; 569164BD CRC32;

Query Match      82.4%; Score 28; DB 1; Length 173;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKKT 7
DB 40 KFLKTI 46
   ||:||||
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:33 ; Search time 148.39 Seconds  
(without alignments)  
2.903 Million cell updates/sec

Title: US-09-124-280A-6

Perfect score: 34

Sequence: 1 KFLKRTL 7

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPREMBL\_10.\*

- 1: sp-archaea.\*
- 2: sp-bacteria.\*
- 3: sp-fungi.\*
- 4: sp-human.\*
- 5: sp-invertebrate.\*
- 6: sp-mammal.\*
- 7: sp-mhc.\*
- 8: sp-organelle.\*
- 9: sp-phage.\*
- 10: sp-plant.\*
- 11: sp-rodent.\*
- 12: sp-virus.\*
- 13: sp-vertebrate.\*
- 14: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	743	4	O43373 homo sapien
2	31	91.2	189	10	O49581 arabidopsis
3	31	91.2	178	10	O92028 arabidopsis
4	30	88.2	305	2	O92028 arabidopsis
5	30	88.2	636	8	O92028 arabidopsis
6	30	88.2	764	10	O64607 arabidopsis
7	30	88.2	544	13	O57325 ambystoma t
8	29	85.3	141	1	O29543 archaeglob
9	29	85.3	322	2	O51656 borrelia bu
10	29	85.3	326	5	P91270 caenorhabdi
11	29	85.3	218	5	O34562 mytilus edu
12	29	85.3	51	8	O34709 hormaphis c
13	29	85.3	114	12	O39830 faba bean n
14	29	85.3	4868	13	O91319 rana catesb
15	28	82.4	396	2	O30714 helicobacte
16	28	82.4	463	2	O85337 escherichia
17	28	82.4	234	2	O92J89 helicobacte
18	28	82.4	537	3	O59831 schizosacch
19	28	82.4	1576	3	O13308 candida alb
20	28	82.4	1576	3	O74209 candida alb
21	28	82.4	2024	4	O15154 homo sapien
22	28	82.4	1792	5	O18983 caenorhabdi
23	28	82.4	622	5	O21389 caenorhabdi
24	28	82.4	349	5	O45849 caenorhabdi
25	28	82.4	4151	5	O96936 drosophila
26	28	82.4	4859	6	O95201 mustela vis
27	28	82.4	609	10	O920A8 arabidopsis
28	28	82.4	1064	11	O60591 mus musculus
29	28	82.4	4869	13	O90985 gallus galli

30	27	79.4	433	2	P71096 bacteroides
31	27	79.4	293	2	O50644 lactococcus
32	27	79.4	249	2	O66909 aquifex aeo
33	27	79.4	642	2	O66684 aquifex aeo
34	27	79.4	1596	2	O59167 acetobacter
35	27	79.4	192	2	O44599 buchera ap
36	27	79.4	503	2	O30417 lactococcus
37	27	79.4	202	2	O84499 chlamydia t
38	27	79.4	476	2	O921Y3 borrelia bu
39	27	79.4	1435	3	O03291 saccharomyc
40	27	79.4	313	3	P78873 schizosacch
41	27	79.4	670	3	O42667 schizosacch
42	27	79.4	971	4	O75432 homo sapien
43	27	79.4	791	5	O19148 caenorhabdi
44	27	79.4	761	5	O16947 caenorhabdi
45	27	79.4	173	12	O88184 san miguel

## ALIGNMENTS

### RESULT 1

O43373  
ID O43373 PRELIMINARY; PRT; 743 AA.  
AC O43373;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PAC CLONE DJ0897G10 FROM 7011.23-Q21.1, COMPLETE SEQUENCE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KALICKI J., ELLIOTT G.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA WATERSTON R.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC004082; AAB97937.1; -  
DR PFAM; PF00168; C2; 2.  
SQ SEQUENCE 743 AA; 81865 MW; 4A4DB2BF CRC32;

Query Match 94.1%; Score 32; DB 4; Length 743;  
Best Local Similarity 85.7%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKRTL 7  
||:|||||  
Db 706 KFMKRTL 712

### RESULT 2

O49581  
ID O49581 PRELIMINARY; PRT; 189 AA.  
AC O49581;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DE 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
DE AUXIN INDUCED LIKE-PROTEIN.  
GN F8F16.140  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurossids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BEVAN M., BRANDT P., DOSE S., JARKE D., SCHARFE M., SCHON O.,  
RA HOEISEL J., MEWES H.W., MAYER K., SCHUELLER C.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU ARABIDOPSIS SEQUENCING PROJECT;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021633; CAA16535.1; -  
 SQ SEQUENCE 189 AA; 21417 MW; 83BD79C8 CRC32;

Query Match 91.2%; Score 31; DB 10; Length 189;  
 Best Local Similarity 85.7%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKRTL 7  
 Db 66 KFLKRTL 72

RESULT 3  
 Q9ZQ28 PRELIMINARY; PRT; 178 AA.

AC Q9ZQ28; 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE PUTATIVE AUXIN-INDUCED PROTEIN.  
 GN T28124.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA LIN X., KAUL S., SHEA T.P., FUJII C.Y., SHEN M., VANAKEN S.E.,  
 RA BARNSTEAD M.E., MASON T.M., BORMAN C.L., RONNING C.M., BENITO M.,  
 RA CARRERA A.J., CREASY T.H., BUELL C.R., TOWN C.D., NIERMAN W.C.,  
 RA FRASER C.M., VENTER J.C.;  
 RT "Arabidopsis thaliana chromosome II BAC T28124 genomic sequence."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006403; AAD18112.1; -  
 SQ SEQUENCE 178 AA; 20054 MW; A007D5E4 CRC32;

Query Match 91.2%; Score 31; DB 10; Length 178;  
 Best Local Similarity 85.7%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKRTL 7  
 Db 50 KFLKRTL 56

RESULT 4  
 Q9ZCE0 PRELIMINARY; PRT; 305 AA.

AC Q9ZCE0; 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE PROBABLE INTEGRASE/RECOMBINASE RIPX (XERC).  
 GN RP817.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RA ANDERSSON S.G.E., ZOMORODIPOUR A., ANDERSSON J.O.,  
 RA SICHERTZ-PONTEN T., ALSMARK U.C.M., PODOWSKI R.M., NAEGLUND A.K.,  
 RA ERIKSSON A.S., WINKLER H.H., KURLAND C.G.;  
 RT "The Genome Sequence of Rickettsia prowazekii and the Origin of Mitochondria."

RL Nature 396:133-140(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MADRID E;  
 RA ANDERSSON S.G.E.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ235273; CAA15242.1; -  
 SQ SEQUENCE 305 AA; 35531 MW; F40AD43C CRC32;

Query Match 88.2%; Score 30; DB 2; Length 305;  
 Best Local Similarity 85.7%; Pred. No. 73;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKRTL 7  
 Db 92 KFLKRTL 98

RESULT 5  
 O99479 PRELIMINARY; PRT; 636 AA.

AC O99479; 01-MAY-1999 (TREMELrel. 10, Created)  
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
 DE REVERSE TRANSCRIPTASE HOMOLOG.  
 OS Pavlova lutheri (Monochrysis lutheri).  
 OG Mitochondrion.  
 OC Eukaryota; Haptophyceae; Pavlovales; Pavlova.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCMP 1325;  
 RA EHARA M., INAGAKI Y., WATANABE K.I., OHAMA T.;  
 RT "Group II intron containing a complete multidomain ORF in the haptophyte alga Pavlova lutheri."  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF045691; AAD12069.1; -  
 KW RNA-directed DNA polymerase; Mitochondrion.  
 SQ SEQUENCE 636 AA; 71454 MW; F0A891A3 CRC32;

Query Match 88.2%; Score 30; DB 8; Length 636;  
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKRTL 7  
 Db 357 KFLKRTL 363

RESULT 6  
 O64607 PRELIMINARY; PRT; 764 AA.

AC O64607; 01-AUG-1998 (TREMELrel. 07, Created)  
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TREMELrel. 07, Last annotation update)  
 DE F1707.17 PROTEIN.  
 GN F1707.17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA VISOTSKAIA V.S., SCHWARTZ J.R., TORIUMI M., YU G., KWAN A., OJI O.,  
 RA LIU S., LI J., ARAUJO R., AU M., BRUNDEL V., BUEHLER E., CONWAY A.B.,  
 RA CONWAY A.R., DEWAR K., FENG J., KIM C., KURTZ D., LI Y., PALM C.J.,  
 RA SHINN P., SON H., DAVIS R.W., ECKER J.R., FEDERSPIEL N.A.,  
 RA THEOLOGIS A.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA THEOLOGIS A.;  
 RN Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA THEOLOGIS A.;  
 RN Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA THEOLOGIS;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC003671; AAC18809.1; -;  
 SQ SEQUENCE 764 AA; 85199 MW; 8D4C0AD2 CRC32;

Query Match 88.2%; Score 30; DB 10; Length 764;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 6  
 DB 243 KFLKKT 248  
 |||||

RESULT 7  
 OS 057325 PRELIMINARY; PRT; 544 AA.  
 AC 057325;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE GLUTAMATE TRANSPORTER 5B.  
 GN SEAT5B.  
 OS Ambystoma tigrinum (Tiger salamander).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Caudata; Salamandroides; Ambystomatidae; Ambystoma.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ELIASOF S.; ARRIZA J.L.; LEIGHTON B.H.; KAVANAUGH M.P.; AMARA S.G.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF018260; AAB88290.1; -;  
 DR PFW; PF00375; SDF; 2.  
 DR PROSITE; PS00713; NA\_DICARBOXYL\_SYM\_1; 1.  
 DR PROSITE; PS00714; NA\_DICARBOXYL\_SYM\_2; 1.  
 SQ SEQUENCE 544 AA; 59609 MW; 9B6084F9 CRC32;

Query Match 88.2%; Score 30; DB 13; Length 544;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKKT 7  
 DB 186 KFLKSL 192  
 |||||

RESULT 8  
 OS 029543 PRELIMINARY; PRT; 141 AA.  
 AC 029543;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE POTASSIUM CHANNEL, PUTATIVE.

GN AF0715.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 RC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE; 98049343.  
 RA KLENK H.-P.; CLAYTON R.A.; TOMB J.-F.; WHITE O.; NELSON K.E.;  
 RA KETCHUM K.A.; DODSON R.J.; GWINN M.; HICKEY E.K.; PETERSON J.D.;  
 RA RICHARDSON D.L.; KERLAVAGE A.R.; GRAHAM D.E.; KYRIDES N.C.;  
 RA FLEISCHMANN R.D.; QUACKENBUSH J.; LEE N.H.; SUTTON G.G.; GILL S.;  
 RA KIRKNESS E.F.; DOUGHERTY B.A.; MCKENNEY K.; ADAMS M.D.; LOFTUS B.;  
 RA PETERSON S.; REICH C.I.; MCNEIL L.K.; BADGER J.H.; GLODEK A.; ZHOU L.;  
 RA OVERBEK R.; COCAYNE J.D.; WEIDMAN J.F.; McDONALD L.; UTTERBACK T.;  
 RA COTTON M.D.; SPRIGGS T.; ARTIACH P.; KAINE B.P.; SYKES S.M.;  
 RA SADOW P.W.; D'ANDREA K.P.; BOWMAN C.; FUJII C.; GARLAND S.A.;  
 RA MASON T.M.; OLSEN G.J.; FRASER C.M.; SMITH H.O.; WOESE C.R.;  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001055; AAB90525.1; -;  
 DR TIGR; AF0715; -;  
 KW Hypothetical protein; Ionic channel.  
 SQ SEQUENCE 141 AA; 16146 MW; 9E88DFEE CRC32;

Query Match 85.3%; Score 29; DB 1; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FLKKT 7  
 DB 134 FLKKT 139  
 |||||

RESULT 9  
 OS 051656 PRELIMINARY; PRT; 322 AA.  
 AC 051656;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 38.6 KD PROTEIN.  
 GN BB0714.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M.; CASJENS S.; HUANG W.M.; SUTTON G.G.; CLAYTON R.A.;  
 RA LATHIGRA R.; WHITE O.; KETCHUM K.A.; DODSON R.; HICKEY E.K.; GWINN M.;  
 RA DOUGHERTY B.; TOMB J.-F.; FLEISCHMANN R.D.; RICHARDSON D.;  
 RA PETERSON J.; KERLAVAGE A.R.; QUACKENBUSH J.; SALZBERG S.; HANSON M.;  
 RA VAN VUOT R.; PALMER N.; ADAMS M.D.; COCAYNE J.D.; WEIDMAN J.;  
 RA UTTERBACK T.; WATTHEY L.; McDONALD L.; ARTIACH P.; BOWMAN C.;  
 RA GARLAND S.; FUJII C.; COTTON M.D.; HORST K.; ROBERTS K.; HATCH B.;  
 RA SMITH H.O.; VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE001171; AAC67064.1; -;  
 DR TIGR; BB0714; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 322 AA; 38625 MW; 79D82CBA CRC32;

Query Match 85.3%; Score 29; DB 2; Length 322;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLKKTLL 7  
 DB 143 FLKKTLL 149

RESULT 10  
 ID P91270 PRELIMINARY; PRT; 326 AA.  
 AC P91270;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE SIMILAR TO ADP. ATP CARRIER PROTEIN.  
 GN F25B4.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA GATTUNG S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; 064842; AAB37086.1; -.  
 DR PFAM; PF00153; mito\_carr; 1.  
 SQ SEQUENCE 326 AA; 36460 MW; 72063959 CRC32;

Query Match 85.3%; Score 29; DB 5; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKKTLL 7  
 DB 7 FLKKTLL 12

RESULT 11  
 ID Q25462 PRELIMINARY; PRT; 218 AA.  
 AC Q25462;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE VITELLINE COAT LYSIN M7 PRECURSOR.  
 OS Mytilus edulis (Blue mussel).  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;

OC Mytilidae; Mytilus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95155270.  
 RA TAKAGI T., AKIO N., RYUSAKU D., KEI-ICHIRO K.;  
 RT "Isolation, characterization, and primary structure of three major  
 RT proteins obtained from Mytilus edulis sperm.";  
 RL J. Biochem. 116:598-605(1994).  
 DR EMBL; D14731; BAA03551.1; -.  
 DR PFAM; PF00059; lectin\_c; 1.  
 KW Signal.  
 FT SIGNAL 1 55 POTENTIAL.  
 FT CHAIN 39 218 VITELLINE COAT LYSIN M7.  
 SQ SEQUENCE 218 AA; 24354 MW; 9E8B1997 CRC32;

Query Match 85.3%; Score 29; DB 5; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKKTLL 7  
 DB 140 FLKKTLL 145

RESULT 12  
 ID Q34709 PRELIMINARY; PRT; 51 AA.  
 AC Q34709;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT I (FRAGMENT).  
 OS Hormaphis cornu.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea;  
 OC Hormaphididae; Hormaphis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94302038.  
 RA STERN D.L.;  
 RT "A phylogenetic analysis of soldier evolution in the aphid family  
 RT Hormaphididae.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 256:203-209(1994).  
 DR EMBL; L27332; AAA50969.1; -.  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 51 AA; 6136 MW; F1A72853 CRC32;

Query Match 85.3%; Score 29; DB 8; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKKTLL 7  
 DB 14 FLKKTLL 19

RESULT 13  
 ID O39830 PRELIMINARY; PRT; 114 AA.  
 AC O39830;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE PUTATIVE MOVEMENT PROTEIN.  
 GN C4.  
 OS faba bean necrotic yellows virus.  
 OC Viruses; ssDNA viruses; Circoviridae.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-SV292-88;  
 RX MEDLINE; 97360021.  
 RA KATUL L., MAISS E., MOROZOV S., VETTEN H.J.;  
 RT "Analysis of six DNA components of the faba bean necrotic yellows  
 virus genome and their structural affinity to related plant virus  
 genomes.";  
 RL Virology 233:247-259(1997).  
 DR EMBL; Y11407; CAA72211.1; -;  
 SQ SEQUENCE 114 AA; 12882 MW; F9C63E72 GRC32;

Query Match 85.3%; Score 29; DB 12; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKRTL 7  
 |||||  
 Db 54 FLKRTL 59

## RESULT 14

Q91319 PRELIMINARY; PRT; 4868 AA.  
 AC Q91319;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE BETA-RYANODINE BINDING PROTEIN.  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 CC Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Raninae; Rana.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SKELETAL MUSCLE;  
 RX MEDLINE; 94274714.  
 RA OYAMADA H., MURAYAMA T., TAKAGI T., IINO M., IWABE N., MIYATA T.,  
 RA OGAWA Y., ENDO M.;  
 RT "Primary structure and distribution of ryanodine-binding protein  
 RT isoforms of the bullfrog skeletal muscle.";  
 RL J. Biol. Chem. 269:17206-17214(1994).  
 DR EMBL; D21071; BAA04647.1; -;  
 DR PFAM; PF00036; ehand; 1.  
 DR PFAM; PF01365; RYDR\_ITPR; 1.  
 DR PFAM; PF00622; SPRY; 3.  
 SQ SEQUENCE 4868 AA; 553053 MW; 99A83E7C CRC32;

Query Match 85.3%; Score 29; DB 13; Length 4868;  
 Best Local Similarity 85.7%; Pred. No. 13e+03;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKRTL 7  
 |||||  
 Db 2822 KFLKRTL 2828

## RESULT 15

O30714 PRELIMINARY; PRT; 396 AA.  
 AC O30714;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
 DE PYRUVATE:FERREDOXIN OXIDOREDUCTASE ALPHA SUBUNIT.  
 OS Helicobacter pylori (Campylobacter pylori).  
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 CC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HP439;  
 RA HAN H., GOODWIN A., HOFFMAN P.S.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN-HP439;  
 RA HAN H., GOODWIN A., HOFFMAN P.S.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF013980; AAC02293.1; -;  
 KW Pyruvate.  
 SQ SEQUENCE 396 AA; 43610 MW; E5AFDFF9 CRC32;

Query Match 82.4%; Score 28; DB 2; Length 396;  
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKRTL 7  
 |||||  
 Db 375 KFLKKS 381

Search completed: September 7, 1999, 20:34:36  
 Job time: 19732 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:09 ; Search time 147.16 Seconds  
(without alignments)  
1.127 Million cell updates/sec

Title: US-09-124-280A-7

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	32	1	R33514
2	35	100.0	11	1	R33532
3	35	100.0	7	1	R33533
4	35	100.0	11	1	R33294
5	35	100.0	7	1	R39295
6	35	100.0	7	1	R711778
7	35	100.0	10	1	R711779
8	35	100.0	12	1	R711791
9	35	100.0	11	1	W21625
10	35	100.0	7	1	W21626
11	35	100.0	10	1	W21629
12	35	100.0	7	1	W21595
13	35	100.0	10	1	W21596
14	35	100.0	12	1	W21608
15	30	85.7	6	1	R33515
16	29	82.9	1279	1	W98761
17	28	80.0	887	1	R65492
18	28	80.0	116	1	W52989
19	27	77.1	405	1	R05713
20	27	77.1	404	1	R48674
21	27	77.1	869	1	R53732
22	27	77.1	10	1	R711781
23	27	77.1	348	1	R85333
24	27	77.1	347	1	R85334
25	27	77.1	2165	1	R94930
26	27	77.1	420	1	W10168
27	27	77.1	9	1	W21698
28	27	77.1	10	1	W21598
29	27	77.1	348	1	W15093
30	27	77.1	347	1	W15094
31	27	77.1	1580	1	W57412
32	27	77.1	2166	1	W48715
33	27	77.1	2166	1	W48716
34	27	77.1	2166	1	W48717
35	27	77.1	2166	1	W48718
36	27	77.1	2166	1	W48719
37	27	77.1	2166	1	W48714
38	27	77.1	2165	1	W77015
39	27	77.1	500	1	W93330
40	27	77.1	107	1	Y11315
41	27	77.1	393	1	Y11025
42	27	77.1	1522	1	Y03183
43	27	77.1	2166	1	Y02005

#### ALIGNMENTS

##### RESULT 1

R33514  
ID R33514 standard; peptide; 32 AA.  
AC R33514:  
DT 07-JUL-1993 (first entry)  
DE Generic peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 2; Page 32; 39pp; English.  
CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X where n = 1-10, the peptide is useful for treating or preventing septic shock, mixing with polymyxin B to reduce its toxicity;  
CC septic shock, mixing with polymyxin B to reduce its toxicity;  
CC removing endotoxins from blood, sera or other fluids (in vivo or in vitro); controlling release of cytokines induced by endotoxins; as diagnostic reagents to detect and quantify toxins in blood or sera; preparing non-toxic antigenic complexes of lipid A or lipopolysaccharide (LPS); and for treating pertussis, bacterial meningitis and HIV-related infections. The usual dose is 10-100 ug/kg/day, given parenterally. It binds to the same sites as polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It has no antibiotic activity; does not lyse erythrocytes; has no toxicity in mice when injected at 50mg/kg and is relatively unstable against proteases.  
SQ Sequence 32 AA;

Query Match 100.0%; Score 35; DB 1; Length 32;

Best Local Similarity 100.0%; Pred. No. 0.66;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

DB 2 KFLKFLK 8

##### RESULT 2

R33532  
ID R33532 standard; peptide; 11 AA.

AC R33532:

DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;

KW cytokine release control; treatment; pertussis; bacterial meningitis;

KW HIV related infections; polymyxin B.

OS Synthetic.

PN ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M.

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated  
 PS Claim 10: Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKFLK 7  
 DB 2 KFLKFLK 8

RESULT 3  
 ID R33533  
 AC R33533; standard; peptide; 7 AA.  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZAG200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 11: Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKFLK 7  
 DB 1 KFLKFLK 7

RESULT 4  
 ID R39294  
 AC R39294; standard; peptide; 11 AA.  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKFLK 7  
 DB 2 KFLKFLK 8

RESULT 5  
 ID R39295  
 AC R39295; standard; peptide; 7 AA.  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent

PT septic shock, in vaccines to detoxify blood, etc.  
PS Claim 11; Page 33; 49pp; English.  
CC The sequence is that of a peptide which binds to the lipid A  
CC component of bacterial endotoxin at the same site as polymyxin B  
CC (PMB) and with about the same affinity to produce a non-toxic  
CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
CC proteolytic degradation in serum, has no antibiotic activity and no  
CC haemolytic action. It is especially used to treat or prevent septic  
CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
CC sera, vaccines, drug solns. etc.; to control release of cytokines  
CC induced by endotoxins; for in vivo or in vitro detoxification of  
CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
CC products. 7 AA;  
SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

RESULT 6  
R71778  
ID R71778 standard; peptide; 7 AA.  
AC R71778;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 9; Page 20; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 10 AA;  
SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | | |  
Db 1 KFLKFLK 7

RESULT 8  
R71791  
ID R71791 standard; peptide; 12 AA.  
AC R71791;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 22; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
CC include units of formula: (A)n, where A is the cationic amino acid Lys  
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 12 AA;  
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

RESULT 7  
R71779  
ID R71779 standard; peptide; 10 AA.  
AC R71779;  
DT 01-OCT-1995 (first entry)

DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 10; Page 20; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 10 AA;  
SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | | |  
Db 1 KFLKFLK 7

RESULT 8  
R71791  
ID R71791 standard; peptide; 12 AA.  
AC R71791;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 22; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
CC include units of formula: (A)n, where A is the cationic amino acid Lys  
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
CC Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 12 AA;  
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 DB 2 KFLKFLK 8

## RESULT 9

W21625  
 ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 DB 2 KFLKFLK 8

## RESULT 10

W21626  
 ID W21626 standard; peptide; 7 AA.  
 AC W21626;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #38.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 42; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 DB 1 KFLKFLK 7

## RESULT 11

W21629  
 ID W21629 standard; peptide; 10 AA.  
 AC W21629;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #41.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 DB 2 KFLKFLK 8

## RESULT 12

W21595  
 ID W21595 standard; peptide; 7 AA.  
 AC W21595;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #7.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 11; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

RESULT 13  
W21596  
ID W21596 standard; peptide; 10 AA.  
AC W21596;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #8.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 03-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI: 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 12; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

RESULT 14  
W21608  
ID W21608 standard; peptide; 12 AA.  
AC W21608;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #20.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.

FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PI (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI: 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 24; Page 26; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 2 KFLKFLK 8

RESULT 15  
R33515  
ID R33515 standard; peptide; 6 AA.  
AC R33515;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B; Group I.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT region 1..3  
FT region /note- "repeat region"  
FT region 4..6  
FT region /note- "repeat region"  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (FORR/) PORRO M.  
PI Porro M;  
DR WPI: 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Example; Page 5; 39pp; English.  
CC The (Group I) peptide is an example of a generic peptide of formula  
CC R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where n = 1-100  
CC and each R is H, an amino acid residue or a fatty acid residue.  
CC The peptide is useful for treating or preventing septic shock,  
CC mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 6 AA;

Query Match 85.7%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKEL 6  
      |||||  
Db 1 KFLKEL 6

Search completed: September 7, 1999, 20:37:10  
Job time: 18459 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:33 ; Search time 80.79 Seconds  
(without alignments)  
0.855 Million cell updates/sec

Title: US-09-124-280A-7

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	11	1	US-08-049-871-8
2	35	100.0	11	1	US-07-819-893-8
3	35	100.0	7	1	US-07-819-893-9
4	35	100.0	11	1	US-08-280-397-8
5	35	100.0	7	1	US-08-280-397-9
6	35	100.0	7	1	US-08-097-830E-7
7	35	100.0	10	1	US-08-097-830E-8
8	35	100.0	12	1	US-08-097-830E-20
9	35	100.0	7	2	US-08-456-112B-7
10	35	100.0	10	2	US-08-456-112B-8
11	35	100.0	12	2	US-08-456-112B-20
12	35	100.0	11	2	US-08-456-112B-37
13	35	100.0	10	2	US-08-456-112B-38
14	35	100.0	10	2	US-08-456-112B-41
15	28	80.0	887	1	US-08-215-709-1
16	28	80.0	116	1	US-08-702-344-21
17	27	77.1	10	1	US-08-097-830E-10
18	27	77.1	2165	2	US-08-514-975B-2
19	27	77.1	10	2	US-08-456-112B-10
20	27	77.1	9	2	US-08-456-112B-40
21	27	77.1	2165	3	PCT-US95-12507-2
22	27	77.1	420	3	PCT-US96-08295-40
23	26	74.3	13	1	US-08-465-325-135
24	26	74.3	773	2	US-08-456-325-135
25	26	74.3	772	2	US-08-524-757-6
26	26	74.3	185	2	US-08-338-343-2
27	26	74.3	331	2	US-08-878-989-21
28	25	71.4	21	1	US-07-908-455A-63
29	25	71.4	287	1	US-08-457-245-3
30	25	71.4	21	1	US-08-434-120-89
31	25	71.4	21	1	US-08-465-325-87
32	25	71.4	10	1	US-08-465-325-121
33	25	71.4	11	1	US-08-465-325-122
34	25	71.4	14	1	US-08-465-325-125
35	25	71.4	527	2	US-08-246-583-2
36	25	71.4	737	2	US-08-185-432-2
37	25	71.4	737	2	US-08-185-432-4
38	25	71.4	224	2	US-08-185-432-7
39	25	71.4	7	2	US-08-944-133-9

ALIGNMENTS

RESULT 1

US-08-049-871-8

; Sequence 8, Application US/08049871

; Patent No. 5358933

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the

; TITLE OF INVENTION: Prevention and Treatment of Septic

; TITLE OF INVENTION: Shock

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson, Costigan & Hoare

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/049,871

; FILING DATE:

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/658,744

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8998

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

US-08-049-871-8

Query Match

Best Local Similarity 100.0%; Score 35; DB 1; Length 11;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KFLKFLK 7

Db 2 KFLKFLK 8

RESULT 2

US-07-819-893-8

; Sequence 8, Application US/07819893

; Patent No. 5371186

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Synthetic Peptides for Detoxification

; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
US-07-819-893-8

Query Match 100.0%; Score 35; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
DB 2 KFLKFLK 8

RESULT 3  
US-07-819-893-9  
;; Sequence 9, Application US/07819893  
;; Patent No. 5371186  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
US-07-819-893-9

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
DB 1 KFLKFLK 7

RESULT 4  
US-08-280-397-8  
;; Sequence 8, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-280-397-8

Query Match 100.0%; Score 35; DB 1; Length 11;



Best Local Similarity 100.0%; Pred. No. 0.38; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 2 KFLKFLK 8

## RESULT 5

US-08-280-397-9  
; Sequence 9, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/280,397

FILING DATE: 07/26/94

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/819,893

FILING DATE: 01/16/92

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-002A

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-280-397-9

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e-04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

## RESULT 6

US-08-097-830E-7  
; Sequence 7, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.

; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/097,830E

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-003

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-097-830E-7

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e-04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
| | | | |  
Db 1 KFLKFLK 7

## RESULT 7

US-08-097-830E-8  
; Sequence 8, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/097,830E

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-003

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-8

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 8  
US-08-097-830E-20  
Sequence 20, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:  
OTHER INFORMATION: sulfide bond between Cys and Cys

Query Match 100.0%; Score 35; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 9  
US-08-456-112B-7  
Sequence 7, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-7

Query Match 100.0%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 10  
US-08-456-112B-8  
Sequence 8, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,112B  
 FILING DATE: May 31, 1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Costigan, James V.  
 REGISTRATION NUMBER: 25,669  
 REFERENCE/DOCKET NUMBER: 576-004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 302-8989  
 TELEFAX: (212) 302-8998  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10 amino acids  
 TYPE: amino acid  
 TOPOLOGY: circular  
 US-08-456-112B-8

Query Match 100.0%; Score 35; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 Db 1 KFLKFLK 7

RESULT 11  
 US-08-456-112B-20  
 ; Sequence 20, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: POIRO, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 12 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 US-08-456-112B-20

Query Match 100.0%; Score 35; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.41;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 Db 2 KFLKFLK 8

RESULT 12  
 US-08-456-112B-37  
 ; Sequence 37, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: POIRO, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 37:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 US-08-456-112B-37

Query Match 100.0%; Score 35; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 |||||  
 Db 2 KFLKFLK 8

RESULT 13  
 US-08-456-112B-38  
 ; Sequence 38, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: POIRO, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA: US/08/456.112B  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-38

Query Match 100.0%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
Db 1 KFLKFLK 7

RESULT 14  
US-08-456-112B-41  
Sequence 41, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porto, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-41

Query Match 100.0%; Score 35; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
Db 2 KFLKFLK 8

RESULT 15  
US-08-215-709-1  
Sequence 1, Application US/08215709  
Patent No. 5432071  
GENERAL INFORMATION:  
APPLICANT: ICHIKAWA, Toshio  
APPLICANT: KOYAMA, Yasuji  
APPLICANT: OTAKE, Hideko  
APPLICANT: NAKANO, Eiichi  
TITLE OF INVENTION: Variant E1 Protein Gene For Pyruvate  
TITLE OF INVENTION: Dehydrogenase Complex And Variant E1 Protein Of Pyruvate  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/215,709  
FILING DATE: 22-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jean B. Fordis  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04853.0011-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 887 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-215-709-1

Query Match 80.0%; Score 28; DB 1; Length 887;  
Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
Db 207 KFLKYLE 213

Search completed: September 7, 1999, 22:38:33

Job time: 7917 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:21 ; Search time 116.8 Seconds  
(without alignments)  
2.401 Million cell updates/sec

Title: US-09-124-280A-7  
Perfect score: 35  
Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 122810 seqs, 40063486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	91.4	236	A56509	peroxisomal membra
2	31	88.6	395	C71148	hypothetical prote
3	30	85.7	956	A45506	H+-transporting AT
4	30	85.7	956	A43637	H+-transporting AT
5	30	85.7	956	S50751	H+-transporting AT
6	30	85.7	951	S52728	H+-transporting AT
7	30	85.7	318	D64417	transformation sen
8	30	85.7	260	A69058	2-hydroxyhepta-2,4
9	30	85.7	141	C71702	hypothetical prote
10	30	85.7	949	T02083	H+-transporting AT
11	30	85.7	394	S64505	HGH1 protein - yea
12	30	85.7	309	S61665	probable membrane
13	30	85.7	1817	D71606	hypothetical prote
14	30	85.7	493	B70417	conserved hypothet
15	30	85.7	141	B70485	hypothetical prote
16	29	82.9	370	H70423	oxygen-independent
17	29	82.9	564	S50520	hexose transport p
18	29	82.9	564	S63405	hexose transport p
19	29	82.9	312	C71146	hypothetical prote
20	29	82.9	569	A45624	trophozoite cystel
21	29	82.9	2291	S11238	polymerase - Berne
22	29	82.9	1279	E54709	type IIS restricti
23	29	82.9	121	C71696	hypothetical prote
24	29	82.9	475	H71660	nitrogen assimilat
25	29	82.9	305	S77033	hypothetical prote
26	29	82.9	286	B70614	probable mmaAl pro
27	29	82.9	287	B64201	hypothetical prote
28	29	82.9	346	I64237	hypothetical prote
29	29	82.9	140	H64411	hypothetical prote
30	29	82.9	337	S53846	ribosomal protein
31	28	80.0	887	D5ECV9	pyruvate dehydroge
32	28	80.0	201	S62757	NADH dehydrogenase
33	28	80.0	718	C73088	glycogen operon pr
34	28	80.0	360	C70185	carotenoid biosynt
35	28	80.0	448	B69745	phosphoglucomutase
36	28	80.0	235	B70009	two-component resp
37	28	80.0	317	D64233	hypothetical prote
38	28	80.0	558	D538614	hypothetical prote
39	28	80.0	2167	S50658	bud emergence prot

ALIGNMENTS

RESULT 1

A56509  
peroxisomal membrane protein PMP27 - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein AOE236; protein O0454; protein YOL147c  
C;Species: Saccharomyces cerevisiae  
C;Date: 21-Jul-1995 #sequence.revision 28-Jul-1995 #text.change 05-Dec-1997  
C;Accession: A56509; B56509; S60389; S62134; S62148; S66844; S57677  
R;Erdmann, R.; Biobel, G.  
J. Cell Biol. 128, 509-523, 1995  
A;Title: Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae lacking the  
A;Reference number: A56509; MUID:95164555  
A;Accession: A56509  
A;Molecule type: DNA  
A;Residues: 1-236 <ERD>  
A;Cross-references: GB:X81465; NID:g683700; PID:g683701  
A;Accession: B56509  
A;Molecule type: protein  
A;Residues: 2-28 <ER2>  
R;Casamayor, A.; Aldea, M.; Casas, C.; Herrero, E.; Gamo, F.J.; Lafuente, M.J.; Gance  
Yeast 11, 1281-1288, 1995  
A;Title: DNA sequence analysis of a 13 kbp fragment of the left arm of yeast chromoso  
A;Reference number: S60385; MUID:96132030  
A;Accession: S60389  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-236 <CAW>  
A;Cross-references: EMBL:Z48239; NID:gl163073; PID:g886949  
R;Marshall, P.A.; Krimkevich, Y.I.; Lark, R.H.; Dyer, J.M.; Veenhuis, M.; Goodman, J.  
J. Cell Biol. 129, 345-355, 1995  
A;Title: Pmp27 promotes peroxisomal proliferation.  
A;Reference number: S62134  
A;Accession: S62134  
A;Molecule type: DNA  
A;Residues: 1-236 <MAR>  
A;Cross-references: EMBL:Z46846; NID:g791131; PID:g791132  
A;Accession: S62148  
A;Molecule type: protein  
A;Residues: 2-30:192-207 <MAW>  
R;Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas  
submitted to the Protein Sequence Database, July 1996  
A;Reference number: S66814  
A;Accession: S66844  
A;Molecule type: DNA  
A;Residues: 1-236 <ARI>  
A;Cross-references: EMBL:Z74889; NID:gl420049; PID:e251925; PID:gl420050; MIPS:YOL147  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:PEX11; PMP27  
A;Cross-references: MIPS:YOL147c; SGD:S0005507  
A;Map position: 15L  
C;Function:  
A;Description: promotion of peroxisomal proliferation  
C;Superfamily: peroxisomal membrane protein PMP27  
C;Keywords: peroxisome; transmembrane protein  
F:2-236/Product: peroxisomal membrane protein PMP27 #status experimental <MAT>

Query Match 91.4% Score 32; DB 2; Length 236;  
Best Local Similarity 85.7% Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 KFLKFLK 7

Db 60 KFLRFLK 66  
|||||

## RESULT 2

hypothetical protein PH0397 - Pyrococcus horikoshii  
C71148  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998  
C:Accession: C71148  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A:Reference number: A71000; MUID:98344137  
A:Accession: C71148  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-395 <KAW>  
A:Cross-references: GB:AF000002; NID:g3236129; PID:d1030415; PID:g3256789  
A:Experimental source: strain OR3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0397

Query Match 88.6%; Score 31; DB 2; Length 395;  
Best Local Similarity 85.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
|||||

Db 237 KFLRFLK 243

## RESULT 3

A45506  
H+-transporting ATPase (EC 3.6.1.35) LHA1 - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998  
C:Accession: A45506  
R:Ewing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.  
Plant Physiol. 94, 1874-1881, 1990  
A:Title: Molecular cloning of tomato plasma membrane H<sup>+</sup>-ATPase.  
A:Reference number: A45506  
A:Accession: A45506  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-956 <EMI>  
A:Cross-references: GB:M60166; NID:g170463; PID:g170464  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (Lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
|||||

Db 61 KFLKFL 66

## RESULT 4

A43637  
H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco  
C:Species: Nicotiana glauca (curled-leaved tobacco)  
C>Date: 11-Feb-1993 #sequence\_revision 12-Mar-1993 #text\_change 17-Mar-1999  
C:Accession: A43637  
R:Boutry, M.; Michelet, B.; Goffeau, A.

Biochem. Biophys. Res. Commun. 162, 567-574, 1989  
A:Title: Molecular cloning of a family of plant genes encoding a protein homologous t  
A:Reference number: A43637; MUID:89334860  
A:Accession: A43637  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-956 <BOU>  
A:Cross-references: GB:M27888; NID:g170205; PID:g170206  
A:Note: the authors translated the codon CGG for residue 131 as Trp  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding doma  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (Lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
|||||

Db 61 KFLKFL 66

## RESULT 5

S50751  
H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato  
C:Species: Solanum tuberosum (potato)  
C>Date: 14-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 24-Apr-1998  
C:Accession: S50751; S39313  
R:Harms, K.; Woehner, R.V.; Schulz, B.; Frommer, W.B.  
Plant Mol. Biol. 26, 979-988, 1994  
A:Title: Isolation and characterization of P-type H<sup>+</sup>-ATPase genes from potato.  
A:Reference number: S50751  
A:Accession: S50751  
A:Molecule type: mRNA  
A:Residues: 1-956 <HAR>  
A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003  
A:Note: the nucleotide is not given  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding doma  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (Lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
|||||

Db 61 KFLKFL 66

## RESULT 6

S52728  
H+-transporting ATPase (EC 3.6.1.35) - kidney bean  
C:Species: Phaseolus vulgaris (kidney bean)  
C>Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 24-Apr-1998  
C:Accession: S52728  
R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S52728  
A:Accession: S52728  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-951 <CAM>  
A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250  
C:Superfamily: Na<sup>+</sup>/K<sup>+</sup>-transporting ATPase alpha chain; ATPase nucleotide-binding doma  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:483-655/Domain: ATPase nucleotide-binding domain homology <ATN>



F:331/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:425/Binding site: ATP (Lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 951;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
| | | | |  
Db 59 KFLKFL 64

RESULT 7  
D64417  
transformation sensitive protein homolog - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: D64417  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: D64417  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <BUL>  
A:Cross-references: GB:U67537; GB:L77117; NID:gl591605; PID:gl591608; TIGR:MJ0940; PID:g  
C:Genetics:  
C:Gene: FOR869481-870437  
A:Map position: transformation sensitive protein; tetratricopeptide repeat homology  
F:17-50/Domain: tetratricopeptide repeat homology <TT1>  
F:51-83/Domain: tetratricopeptide repeat homology #status atypical <TT2>  
F:84-117/Domain: tetratricopeptide repeat homology <TT3>  
F:118-151/Domain: tetratricopeptide repeat homology <TT4>  
F:152-185/Domain: tetratricopeptide repeat homology <TT5>  
F:186-219/Domain: tetratricopeptide repeat homology <TT6>  
F:255-288/Domain: tetratricopeptide repeat homology <TT7>

Query Match 85.7%; Score 30; DB 2; Length 318;  
Best Local Similarity 85.7%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFL 7  
| | | | |  
Db 51 KFLKFL 57

RESULT 8  
A69068  
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase - Methanobacterium thermoautotrophicum (st  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Sep-1998  
C:Accession: A69068  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Qiu, D.; Spadator, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: A69068  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-260 <MTH>  
A:Cross-references: GB:AE000911; GB:AE000666; NID:g2622623; PID:g2622625  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1507  
A:Start codon: GTG

C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 85.7%; Score 30; DB 2; Length 260;  
Best Local Similarity 71.4%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 7  
| | | | |  
Db 5 KFLKFL 11

RESULT 9  
C71702  
hypothetical protein RP436 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 21-Nov-1998  
C:Accession: C71702  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630  
A:Accession: C71702  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-141 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3860788; PID:el342737; PID:g386099  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: RP436

Query Match 85.7%; Score 30; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
| | | | |  
Db 118 KFLKFL 123

RESULT 10  
T02083  
H+-transporting ATPase (EC 3.6.1.35) Mha1 - maize  
C:Species: Zea mays (maize)  
C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 26-Feb-1999  
C:Accession: T02083  
R:Jin, Y.; Bennetzen, J.L.  
Plant Cell 6, 1177-1186, 1994  
A:Title: Integration and nonrandom mutation of a plasma membrane proton ATPase gene f  
A:Reference number: Z14555  
A:Accession: T02083  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-949 <JIN>  
A:Cross-references: EMBL:U09989; NID:g507770; PID:g533775  
A:Experimental source: strain D3L  
C:Genetics:  
A:Gene: Mha1  
A:Map position: 2L  
A:Introns: 21/3; 61/3; 94/3; 140/3; 185/3; 245/3; 314/3; 354/3; 394/3; 435/3; 470/3;  
C:Function:  
A:Description: transports protons across the plasma membrane to regulate intracellular  
C:Keywords: ATP; hydrogen ion transport; hydrolase; phosphoprotein

Query Match 85.7%; Score 30; DB 2; Length 949;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
| | | | |  
Db 64 KFLKFL 69

```

RESULT 11
S64505
HGHI protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G7538; protein YGR187c
C:Species: Saccharomyces cerevisiae
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 06-Feb-1998
C:Accession: S64505; S59348
R:Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64499
A:Accession: S64505
A:Molecule type: DNA
A:Residues: 1-394 <ARR>
A:Cross-references: EMBL:Z72972; NID:g1323331; PID:e243565; PID:g1323332; MIPS:YGR187c
A:Experimental source: strain S288C
R:Sun, Z.; Liang, J.; Hampsey, M.
submitted to the EMBL Data Library, May 1995
A:Description: The HGHI gene encodes a protein with structural similarity to mammalian H
A:Reference number: S59348
A:Accession: S59348
A:Molecule type: DNA
A:Residues: 111-394 <SUN>
A:Cross-references: EMBL:U27358; NID:g1002515; PID:g1002516
C:Genetics:
A:Gene: SGD:HGHI
A:Cross-references: SGD:S0003419; MIPS:YGR187c
A:Map position: 7R
C:Keywords: transmembrane protein
F:296-312/Domain: transmembrane #status predicted <TMM>

Query Match 85.7%; Score 30; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6
| | | | |
Db 91 KFLKFL 96

RESULT 12
S61665
Probable membrane protein YOR107w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Q3224; hypothetical protein YOR3224w
C:Species: Saccharomyces cerevisiae
C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 21-Nov-1997
C:Accession: S61665; S66992
R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia
submitted to the EMBL Data Library, December 1995
A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
A:Reference number: S61643
A:Accession: S61665
A:Molecule type: DNA
A:Residues: 1-309 <BEN>
A:Cross-references: EMBL:X94335; NID:g1262139; PID:e217829; PID:g1164952
R:Vooss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoerge, W.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66965
A:Accession: S66992
A:Molecule type: DNA
A:Residues: 1-309 <VOS>
A:Cross-references: EMBL:Z75015; NID:g1420290; PID:e252009; PID:g1420291; MIPS:YOR107w
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 15R
C:Keywords: transmembrane protein
F:129-145/Domain: transmembrane #status predicted <TMM>

Query Match 85.7%; Score 30; DB 2; Length 309;
Best Local Similarity 85.7%; Pred. No. 41;

```

```

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7
| | | | |
Db 50 KFYKFLK 56

RESULT 13
D71606
hypothetical protein PFB0800c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: D71606
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: D71606
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1817 <GAR>
A:Cross-references: GB:AE001418; GB:AE001362; NID:g3845275; PID:g3845280; TIGR:PFB080
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0800c

Query Match 85.7%; Score 30; DB 2; Length 1817;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7
| | | | |
Db 810 FLKFLK 815

RESULT 14
B70417
conserved hypothetical protein aq_1350 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
C:Accession: B70417
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: B70417
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <AQF>
A:Cross-references: GB:AE000736; NID:g2983763; PID:g2983778; GB:AE000657
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_1350

Query Match 85.7%; Score 30; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6
| | | | |
Db 188 KFLKFL 193

RESULT 15
B70485
hypothetical protein aq_2157 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
C:Accession: B70485

```

R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: B70485  
A>Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-141 <AAQF>  
A:Cross-references: GB:AE000776; NID:g2984355; PID:g2984366; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_2157

Query Match 85.7%; Score 30; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
| | | | |  
Db 38 KFLKFL 43

Search completed: September 7, 1999, 23:06:22  
Job time: 2470 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:07 ; Search time 71.87 Seconds  
(without alignments)  
2.753 Million cell updates/sec

Title: US-09-124-280A-7

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	91.4	235	1 PEXB_YEAST	Q12462 saccharomyc
2	30	85.7	394	1 HGHI_YEAST	P48362 saccharomyc
3	30	85.7	956	1 PMAI_LYCES	P22180 lycopersico
4	30	85.7	103	1 RR6_CYACA	Q19917 cyanidium c
5	30	85.7	318	1 Y940_METJA	Q58350 methanococc
6	29	82.9	569	1 CYSF_PLAFA	P25805 plasmodium
7	29	82.9	140	1 FLAE_METJA	Q58306 methanococc
8	29	82.9	564	1 HXTC_YEAST	P39924 saccharomyc
9	29	82.9	564	1 HXTC_YEAST	P39924 saccharomyc
10	29	82.9	2291	1 RRPB_BEV	P18458 berne virus
11	29	82.9	337	1 RT07_ACACA	P46756 acanthamoeb
12	29	82.9	287	1 Y011_MYCGE	P47237 mycoplasma
13	29	82.9	346	1 Y343_MYCGE	P47585 mycoplasma
14	28	80.0	2167	1 BEM2_YEAST	P39960 saccharomyc
15	28	80.0	943	1 DIP2_YEAST	Q12220 saccharomyc
16	28	80.0	233	1 DP3E_BUCAP	Q08880 buchnera ap
17	28	80.0	201	1 NU6M_CYACA	P48925 cyanidium c
18	28	80.0	886	1 ODPL_ECOLI	P06938 escherichia
19	28	80.0	578	1 PSP2_YEAST	P50109 saccharomyc
20	28	80.0	599	1 VE1_HPV04	Q07846 human papil
21	28	80.0	317	1 Y302_MYCGE	P47544 mycoplasma
22	28	80.0	558	1 YCX9_ASTLO	P34782 astasia lon
23	28	80.0	628	1 YELF_SCHPO	Q13879 schizosacch
24	28	80.0	823	1 YHX2_YEAST	P38863 saccharomyc
25	27	77.1	279	1 ATPG_MYCGE	P47640 mycoplasma
26	27	77.1	355	1 CAPH_STAAU	P39857 staphylococ
27	27	77.1	626	1 CC23_YEAST	P16522 saccharomyc
28	27	77.1	526	1 CPF5_RAT	P51870 rattus norv
29	27	77.1	817	1 ENPL_CATRO	P35016 catharanthu
30	27	77.1	890	1 IMB2_HUMAN	Q92973 homo sapien
31	27	77.1	451	1 MURD_BORBU	O51532 borrelia bu
32	27	77.1	404	1 NAVA_CLOSO	P15698 clostridium
33	27	77.1	269	1 NIVA_CLOPA	Q00853 clostridium
34	27	77.1	869	1 PLCL_YEAST	P32383 saccharomyc
35	27	77.1	629	1 PR39_YEAST	P39682 saccharomyc
36	27	77.1	469	1 RGS7_BOVIN	O46470 bos taurus
37	27	77.1	420	1 RGS7_HUMAN	P49802 homo sapien
38	27	77.1	469	1 RGS7_MOUSE	O54829 mus musculu
39	27	77.1	179	1 RL15_ARCFU	Q28376 archaeoglob
40	27	77.1	178	1 RL5_ARCFU	Q28367 archaeoglob
41	27	77.1	2238	1 RRPL_BUNIW	P20470 bunyamwera
42	27	77.1	2165	1 RRPL_HRSVA	P28887 human respi
43	27	77.1	195	1 RUBY_CLOPE	P51591 clostridium

ALIGNMENTS

RESULT 1

```
PEXB_YEAST
ID PEXB_YEAST STANDARD; PRT; 235 AA.
AC Q12462;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PEROXISOMAL MEMBRANE PROTEIN PMP27 (PEROXIN-11).
GN PEX11 OR PMP27 OR PMP24 OR YOL147C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.
RC STRAIN=S288C;
RX MEDLINE: 95164555.
RA ERDMANN R., BLOBEL G.;
RT "Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae
RL lacking the peroxisomal membrane protein Pmp27p."
RL J. CELL BIOL. 128:509-523(1995).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-29 AND 191-206.
RC STRAIN=GRF88;
RX MEDLINE: 95238534.
RA MARSHALL P.A., KRIMKEVICH Y.I., LARK R.H., DYER J.M., VEENHUIS M.,
GOODMAN J.M.;
RT "Pmp27 promotes peroxisomal proliferation."
RL J. CELL BIOL. 129:345-355(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE: 96132030.
RA CASAMAYOR A., ALDEA M., CASAS C., HERRERO E., GAMO F.J.,
LAFUENTE M.J., GANCEDO C., ARINO J.;
RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
chromosome XV containing seven new open reading frames."
RL YEAST 11:1281-1288(1995).
CC -1- FUNCTION: INVOLVED IN PEROXISOMAL PROLIFERATION. COULD PARTICIPATE
CC IN PEROXISOMAL ELONGATION OR FISSION. MAY BE INVOLVED IN PARCELING
CC OF PEROXISOMES INTO REGULAR QUANTA.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL MEMBRANE-ASSOCIATED.
CC -1- SIMILARITY: STRONG, TO C.BOIDINII PEX11.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: Z48239; G886949;
CC EMBL: Z74889; E251925;
CC EMBL: X81465; G683701;
CC EMBL: Z46846; G791132;
CC SGD: L0002650; PMP27.
CC PEROXISOME; MEMBRANE.
FT INITMET 0 0
SQ SEQUENCE 235 AA; 26744 MW; 228BDA38 CRC32;
```

Query Match 91.4%; Score 32; DB 1; Length 235;  
Best Local Similarity 85.7%; Pred. No. 6.8;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7

```

Db      59 KFLKFL 65
|||||
RESULT 2
HGHI_YEAST
ID HGHI_YEAST STANDARD; PRT; 394 AA.
AC P48362;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HGHI PROTEIN
GN HGHI OR VGR187C OR G7538.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCETES.
[1]
RN SEQUENCE OF 111-394 FROM N.A.
RA SUN Z., LIANG J., HAMPSEY M.;
RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE: 97279231.
RA ARROYO J., GARCIA-GONZALEZ M., GARCIA-SAEZ M.I., SANCHEZ-PEREZ M.,
RA NOMBELA C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT of Saccharomyces cerevisiae chromosome VII.";
RL YEAST 13:357-363(1997).
CC -1- SIMILARITY: TO S.POMBE SPAC26F1.12C.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U27358; G1002516; -
DR EMBL; 272972; E243565; -
DR EMBL; X99074; E252631; -
DR SGD; L0003003; HGHI.
FT DOMAIN 371 389 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 394 AA; 44951 MW; EE1B34E3 CRC32;

Query Match 85.7%; Score 30; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6
|||||
Db 91 KFLKFL 96

RESULT 3
PMAL_LYCES
ID PMAL_LYCES STANDARD; PRT; 956 AA.
AC P22180;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).
GN LHAI.
OS LYCOPERSICON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPANTAE; EMERYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; SOLANACEAE; SOLANACEAE; SOLANACEAE; SOLANUM.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. CASTELMART; TISSUE=ROOT;
RA EWING N.N., WIMMERS L.E., MEYER D.J., CHETELAT R.T., BENNETT A.B.;

"Molecular cloning of tomato plasma membrane H+-ATPase.";
PLANT PHYSIOL. 94:1874-1881(1990).
-1- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A
HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE
ACTIVE TRANSPORT OF NUTRIENTS BY H+-SYMPTOT. THE RESULTING
EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKALIZATION MAY MEDIATE
GROWTH RESPONSES.
-1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
-1- SUBUNIT: POSSIBLY EXISTS AS AN HOMODIMER OR AN HOMOTRIMER.
-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-1- AS MANY AS 6 TO 8 CLOSELY RELATED GENES MAY ENCODE OTHER ISOFORMS
OF PLASMA MEMBRANE ATPASE IN TOMATO, LIKE THE LHAI GENE PRODUCT
WHICH IS 96% IDENTICAL TO THE LHAI GENE PRODUCT.
-1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
(E1-E2 ATPASES).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M60166; G170464; -
DR PIR; A45506; A45506.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR PFAM; PF00122; E1-E2_ATPase; 1.
KW HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;
KW ATP-BINDING; MULTIGENE FAMILY.
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 1 (POTENTIAL).
FT DOMAIN 86 96 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 97 116 2 (POTENTIAL).
FT DOMAIN 117 245 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 246 267 3 (POTENTIAL).
FT DOMAIN 268 279 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 280 302 4 (POTENTIAL).
FT DOMAIN 303 677 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 678 696 5 (POTENTIAL).
FT DOMAIN 697 712 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 713 736 6 (POTENTIAL).
FT DOMAIN 737 794 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 795 815 7 (POTENTIAL).
FT DOMAIN 816 830 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 831 847 8 (POTENTIAL).
FT DOMAIN 848 956 EXTRACELLULAR (POTENTIAL).
FT MOD_RES 333 333 PHOSPHORYLATION (BY SIMILARITY).
FT BINDING 427 427 ATP (BY SIMILARITY).
SQ SEQUENCE 956 AA; 105103 MW; E798A2DA CRC32;

Query Match 85.7%; Score 30; DB 1; Length 956;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6
|||||
Db 61 KFLKFL 66

RESULT 4
RR6_CYACA
ID RR6_CYACA STANDARD; PRT; 103 AA.
AC O19917;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S6.
GN RPS6.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OC CHLOROPLAST.
OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDIAEAE;

```

```

CC CYANIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK-1;
RA GLOCKNER G., ROSENTHAL A., VALENTIN K.;
RT "Organisation of 46 kb of the Cyanidium caldarium RK1 plasmid genome.";
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: BINDS TOGETHER WITH S18 TO 16S RIBOSOMAL RNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AF022186; G2465743; -.
DR PROSITE; PS01048; RIBOSOMAL_S6; FALSE_NEG.
DR PFAM; PF01250; Ribosomal_S6; 1.
KW RIBOSOMAL PROTEIN; RNA-BINDING; CHLOROPLAST.
SQ SEQUENCE 103 AA; 12340 MW; CC91D65A CRC32;

Query Match 85.7%; Score 30; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6
Db 28 KFLKFL 33

RESULT 5
Y940_METJA STANDARD; PRT; 318 AA.
ID Y940_METJA
AC Q58350;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0940.
GN MJ0940.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
CC
CC SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D., SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D., KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I., OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLOCKER A., SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D., UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C., COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M., KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";
RL SCIENCE 273:1058-1073(1996).
CC -1- SIMILARITY: CONTAINS 2 TPR DOMAINS.
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0941.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

```

```

CC EMBL; U67537; G1591608; -.
DR TIGR; MJ0940; -.
DR PFAM; PF00515; TPR; 2.
KW HYPOTHETICAL PROTEIN; REPEAT; TPR DOMAIN.
SQ SEQUENCE 318 AA; 36861 MW; 7775AE60 CRC32;

Query Match 85.7%; Score 30; DB 1; Length 318;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFL 7
Db 51 KFLKFL 57

RESULT 6
CYSF_PLAFA STANDARD; PRT; 569 AA.
ID CYSF_PLAFA
AC P25805;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE THROPHOZOITE CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-) (TCP).
OS PLASMODIUM FALCIPARUM.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.
CC [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 92228005.
RA ROSENTHAL P.J., NELSON R.G.;
RT "Isolation and characterization of a cysteine proteinase gene of Plasmodium falciparum.";
RL MOL. BIOCHEM. PARASITOL. 51:143-152(1992).
CC -1- FUNCTION: PROBABLY DEGRADES ERYTHROCYTE HEMOGLOBIN.
CC -1- DEVELOPMENTAL STAGE: THROPHOZOITE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE PAPAIN FAMILY OF THIOL PROTEASES.
CC -1- SIMILARITY: STRONGEST SIMILARITY TO CATHEPSIN L.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
CC
CC EMBL; M81341; G160248; -.
DR PIR; A45624; A45624.
DR PROSITE; PS00139; THIOL PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR PFAM; PF00112; Cys-protease; 2.
DR HSP; P07858; IC5B.
KW HYDROLASE; THIOL PROTEASE; ZYMOMEN; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 ?
FT PROPEP 2 332 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 333 569 THROPHOZOITE CYSTEINE PROTEINASE.
FT DOMAIN 64 70 POLY-ASN.
FT ACT_SITE 357 357 BY SIMILARITY.
FT ACT_SITE 488 488 BY SIMILARITY.
FT ACT_SITE 533 533 BY SIMILARITY.
FT DISULFID 354 395 BY SIMILARITY.
FT CARBOHYD 58 58 POTENTIAL.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 121 121 POTENTIAL.
FT CARBOHYD 127 127 POTENTIAL.
FT CARBOHYD 479 479 POTENTIAL.
FT CARBOHYD 487 487 POTENTIAL.
SQ SEQUENCE 569 AA; 66880 MW; 9B8F0096 CRC32;

```

Query Match 82.9%; Score 29; DB 1; Length 569;  
 Best Local Similarity 71.4%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 II IIII I  
 Db 224 KFFKFK 230

RESULT 7  
 ID FLAE\_METJA STANDARD; PRT; 140 AA.  
 AC Q58306.

DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE PUTATIVE FLAGELLA-RELATED PROTEIN E.  
 GN FLAE OR MJ0896  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE: 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KIRKNESS E.F., WEINSTOCK K.G., ADAMS M.D., REICH C.I.,  
 RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., ADAMS M.D., REICH C.I.,  
 RA SCOTT J.L., GEORGHAN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKI M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL SCIENCE 273:1058-1073(1996).  
 CC -!- SIMILARITY: STRONG, TO M.VOLTAE FLAE, ALSO TO FLAD.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 [1]

DR EMBL: U67533; G1591572; -;  
 DR TIGR: MJ0896; -;  
 KW FLAGELLA.  
 SQ SEQUENCE 140 AA; 15891 MW; 61446FEB CRC32;

Query Match 82.9%; Score 29; DB 1; Length 140;  
 Best Local Similarity 85.7%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 I IIIII  
 Db 73 KLLKFLK 79

RESULT 8  
 HXTC\_YEAST STANDARD; PRT; 564 AA.  
 AC P39924;

DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HEXOSE TRANSPORTER HXT13.  
 GN HXT13 OR HXT8 OR YEL069C.  
 OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
 ,OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.

RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA DITRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
 RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,  
 RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNTKE-SMITH S.,  
 RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,  
 RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,  
 RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,  
 RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;  
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 [1]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 [1]

DR EMBL: U18795; G603249; -;  
 DR SGD; L0002640; HXT13.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 RF PFAM: PF00083; sugar\_tr; 1.  
 KW REPEAT: TRANSMEMBRANE; SUGAR TRANSPORT: TRANSPORT; GLYCOPROTEIN.  
 ET DOMAIN 1 51  
 FT TRANSMEM 52 73  
 FT DOMAIN 74 107  
 FT TRANSMEM 108 127  
 FT DOMAIN 128 136  
 FT TRANSMEM 137 156  
 FT DOMAIN 157 163  
 FT TRANSMEM 164 185  
 FT DOMAIN 186 197  
 FT TRANSMEM 198 217  
 FT DOMAIN 218 229  
 FT TRANSMEM 230 250  
 FT DOMAIN 251 318  
 FT TRANSMEM 319 339  
 FT DOMAIN 340 352  
 FT TRANSMEM 353 374  
 FT DOMAIN 375 381  
 FT TRANSMEM 382 401  
 FT DOMAIN 402 421  
 FT TRANSMEM 422 445  
 FT DOMAIN 446 457  
 FT TRANSMEM 458 479  
 FT DOMAIN 480 484  
 FT TRANSMEM 485 506  
 FT DOMAIN 507 564  
 SQ SEQUENCE 564 AA; 62734 MW; 073CAE7E CRC32;

Query Match 82.9%; Score 29; DB 1; Length 564;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 IIIII I  
 Db 555 KFLKFSK 561

RESULT 9  
 HXTG\_YEAST STANDARD; PRT; 564 AA.  
 ID HXTG\_YEAST  
 AC P53631;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)



DE HEXOSE TRANSPORTER HXT17.  
GN HXT17 OR YNR072W OR N3615.  
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCAROMYCETALES;  
OC SACCAROMYCETACEAE; SACCAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDRE B., IRAQUI HOUSAINI I., URRESTARAZU L.A., VISSERS S.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z71687; E239602; -  
CC SGD: L0003265; HXT17.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
DR PFAM; PF00083; sugar\_tr; 1.  
KW REPEAT; TRANSMEMBRANE; SUGAR TRANSPORT; GLYCOPROTEIN.  
FT TRANSMEM 50 70  
FT TRANSMEM 110 130 POTENTIAL.  
FT TRANSMEM 137 157 POTENTIAL.  
FT TRANSMEM 168 188 POTENTIAL.  
FT TRANSMEM 196 216 POTENTIAL.  
FT TRANSMEM 230 250 POTENTIAL.  
FT TRANSMEM 318 338 POTENTIAL.  
FT TRANSMEM 353 373 POTENTIAL.  
FT TRANSMEM 382 402 POTENTIAL.  
FT TRANSMEM 423 443 POTENTIAL.  
FT TRANSMEM 463 483 POTENTIAL.  
FT TRANSMEM 485 505 POTENTIAL.  
SQ SEQUENCE 564 AA; 62828 MW; 23C1F139 CRC32;

Query Match 82.9%; Score 29; DB 1; Length 564;  
Best Local Similarity 85.7%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7  
||| |  
Db 555 KFLKFSK 561

RESULT 10  
RRPB\_BEV  
ID RRPB\_BEV STANDARD; PRT; 2291 AA.  
AC P18458;  
DT 01-NOV-1990 (REL. 16, CREATED)  
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (ORF1B).  
GN POL.  
OS BERNE VIRUS (BEV).  
OS VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; NIDOVIRALES;  
OC CORONAVIRIDAE; TOROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE PI38/72;  
RX MEDLINE; 90356389.  
RA SNIJDER E.J., DEN BOON J.A., BREDBEEK P.J., HORZINEK M.C.,  
RA RIJNBRAND R., SPAAN W.J.M.;  
RT "The carboxyl-terminal part of the putative Berne virus polymerase is  
RT expressed by ribosomal frameshifting and contains sequence motifs  
RT which indicate that toro- and coronaviruses are evolutionarily  
RT related.";

RL NUCLEIC ACIDS RES. 18:4535-4542(1990).  
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE  
CC + RNA(N).  
CC -1- THIS PROTEIN IS EXPRESSED BY AN EFFICIENT RIBOSOMAL FRAMESHIFTING  
CC MECHANISM. RIBOSOMAL FRAMESHIFTING IS AN ELEGANT MECHANISM FOR  
CC REGULATING THE SYNTHESIS OF SEVERAL PROTEINS IN A WELL BALANCED  
CC MANNER.  
CC -1- SIMILARITY: TO CORONAVIRUSES POLYMERASE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X52374; E23105; -  
CC PIR; S11238; S11238.  
KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
FT DOMAIN 515 744 POLYMERASE.  
FT DOMAIN 1099 1374 HELICASE.  
FT NP\_BIND 1101 1108 ATP (BY SIMILARITY).  
FT SIMILAR 845 928 TO ITS HOMOLOG FROM MHV (AA 938-1022)  
FT SIMILAR 1921 2209 AND IBV (AA 929-1013).  
FT SIMILAR TO ITS HOMOLOG FROM MHV (AA 2322-2641)  
FT AND IBV (AA 2274-2598).  
SQ SEQUENCE 2291 AA; 260843 MW; C0283C2D CRC32;

Query Match 82.9%; Score 29; DB 1; Length 2291;  
Best Local Similarity 85.7%; Pred. No. 2,7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7  
||| |  
Db 524 KFSKFLK 530

RESULT 11  
RT07\_ACACA  
ID RT07\_ACACA STANDARD; PRT; 337 AA.  
AC P46756;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.  
GN RPS7.  
OS ACANTHAMOEBA CASTELLANII (AMOEB).  
OG MITOCHONDRION.  
OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 30010 / NEFF;  
RX MEDLINE; 93157849.  
RA LONERGAN K.M., GRAY M.W.;  
RT "Editing of transfer RNAs in Acanthamoeba castellanii mitochondria.";  
RT SCIENCE 259:812-816(1993).  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U12386; G562050; -  
DR PROSITE; PS00052; RIBOSOMAL\_S7; FALSE\_NEG.  
DR PFAM; PF00177; S7; 1.  
KW RIBOSOMAL PROTEIN; MITOCHONDRION.

SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 82.9%; Score 29; DB 1; Length 337;  
Best Local Similarity 71.4%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
DB 328 RELKFLK 334

## RESULT 12

```

ID Y011_MYCGE STANDARD; PRT; 287 AA.
AC P47257; Q49362;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MG011.
GN MG011.
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 158-255 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 158-255 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 94075230.
RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random sequencing."
RL J. BACTERIOL. 175:7918-7930(1993).
CC -!- SIMILARITY: BELONGS TO THE ATP-GRASP SUPPERFAMILY; RIMK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39679; G1045679; -
CC DR EMBL; U02257; G406920; -
CC DR TIGR; MG011; -
CC KW HYPOTHETICAL PROTEIN.
CC RESULT 14
CC BEM2_YEAST STANDARD; PRT; 2167 AA.
CC ID BEM2_YEAST
CC AC P39960;
CC DT 01-FEB-1995 (REL. 31, CREATED)
CC DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
CC DE GTPASE ACTIVATING PROTEIN BEM2/IPL2.
CC GN BEM2 OR IPL2 OR SUP9 OR YER155C.
CC OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
CC OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
CC SACCHAROMYCETACEAE; SACCHAROMYCES.
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN-S288C;
CC RC MEDLINE; 95050996.
CC RX KIM Y., FRANCISCO L., CHEN G., MARCOTTE E., CHAN C.S.;

```

Query Match 82.9%; Score 29; DB 1; Length 287;  
Best Local Similarity 85.7%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
DB 276 KFLKFLK 282

## RESULT 13

```

Y343_MYCGE STANDARD; PRT; 346 AA.
ID Y343_MYCGE
AC P4755;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MG343.
GN MG343.
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 37-150 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 94075230.
RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random sequencing."
RL J. BACTERIOL. 175:7918-7930(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39718; G1046049; -
CC DR EMBL; U01811; G407662; -
CC DR TIGR; MG343; -
CC KW HYPOTHETICAL PROTEIN.
CC SEQUENCE 346 AA; 40834 MW; 1450B76F CRC32;
Query Match 82.9%; Score 29; DB 1; Length 346;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KFLKFLK 7
DB 112 KFLKFLK 118
RESULT 14
BEM2_YEAST STANDARD; PRT; 2167 AA.
ID BEM2_YEAST
AC P39960;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE GTPASE ACTIVATING PROTEIN BEM2/IPL2.
GN BEM2 OR IPL2 OR SUP9 OR YER155C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN-S288C;
CC RC MEDLINE; 95050996.
CC RX KIM Y., FRANCISCO L., CHEN G., MARCOTTE E., CHAN C.S.;

```

RT "Control of cellular morphogenesis by the Ipi2/Bem2 GTPase-activating  
 RL protein: possible role of protein phosphorylation.";  
 J. CELL BIOL. 127:1381-1394(1994).

[2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 95050997.  
 RA PETERSON J., ZHENG Y., BENDER L., MYERS A., CERIONE R., BENDER A.;  
 RT "Interactions between the bud emergence proteins Bem1p and Bem2p and  
 RL rho-type GTPases in yeast.";  
 J. CELL BIOL. 127:1395-1406(1994).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
 RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,  
 RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICK-SMITH S.,  
 RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,  
 RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,  
 RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,  
 RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;  
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [4]  
 RN CHARACTERIZATION.  
 RP MEDLINE: 94123964.  
 RX CHAN C.S., BOTSTEIN D.;  
 RA "Isolation and characterization of chromosome-gain and increase-in-  
 RT ploidy mutants in yeast.";  
 RL GENETICS 135:677-691(1993).  
 CC INVOLVED IN THE CONTROL OF CELLULAR MORPHOGENESIS. REQUIRED FOR  
 CC PROPER BUD SITE SELECTION AND BUD EMERGENCE.  
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 GAP DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: 235159; G511137; -;  
 DR EMBL: L33832; G499695; -;  
 DR EMBL: U18917; G603395; -;  
 DR SGD: L0000168; BEM2.  
 DR PROSITE: PS50003; PH DOMAIN; 1.  
 DR PFAM: PF00617; RasGEF; 1.  
 DR PFAM: PF00618; RasGEF; 1.  
 DR PFAM: PF00620; RhoGAP; 1.  
 DR GTPASE ACTIVATION.  
 KW DOMAIN 16 25 POLY-SER.  
 FT DOMAIN 35 43 POLY-SER.  
 FT DOMAIN 58 63 POLY-HIS.  
 FT DOMAIN 198 208 POLY-ASN.  
 FT DOMAIN 253 260 POLY-SER.  
 FT DOMAIN 1161 1165 POLY-THR.  
 FT DOMAIN 1914 1948 PH.  
 FT DOMAIN 1981 2134 GAP DOMAIN.  
 FT SEQUENCE 2167 AA; 245428 MW; DA2FF455 CRC32;

Query Match 80.0%; Score 28; DB 1; Length 2167;  
 Best Local Similarity 83.3%; Pred. No. 4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKFLK 7  
 |:||||  
 Db 463 FIRFLK 468

RESULT 15  
 DIP2\_YEAST

ID DIP2\_YEAST STANDARD; PRT; 943 AA.  
 AC Q12220; Q05386;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DOM34 INTERACTING PROTEIN 2  
 GN DIP2 OR YLR129W OR L9233.1 OR L3116.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU 2.,  
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
 RA HALLSWORTH K., HAWKINS J., HILLIER L., JER M., JOHNSON D.,  
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
 RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
 RA WILSON R., WATERSTON R.;  
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-S288C;  
 RC DELIUS H.;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [3]  
 RN SEQUENCE OF 1-458 FROM N.A.  
 RC STRAIN-FV23 / RD005;  
 RA VERHASSELT P., VOST M., VOLCKAERT G.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: TO S.POMBE SPBC3D5.12.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U53877; G1256869; -;  
 DR EMBL: U53881; G2580465; -;  
 DR EMBL: X91258; G995691; -;  
 DR EMBL: Z73301; E245568; -;  
 DR EMBL: Z73302; E322129; -;  
 DR EMBL: X89514; E198829; -;  
 DR SGD: L0003088; DIP2.  
 DR PROSITE: PS00678; WD\_REPEATS; 4.  
 DR PFAM: PF00400; G-beta; 8.  
 KW REPEAT; WD REPEAT.  
 FT REPEAT 77 107 WD1.  
 FT REPEAT 119 149 WD2.  
 FT REPEAT 161 190 WD3.  
 FT REPEAT 202 230 WD4.  
 FT REPEAT 389 418 WD5.  
 FT REPEAT 428 458 WD6.  
 FT REPEAT 471 501 WD7.  
 FT REPEAT 571 601 WD8.  
 FT REPEAT 613 643 WD9.  
 FT REPEAT 655 685 WD10.  
 FT SEQUENCE 943 AA; 106342 MW; 7B348EA8 CRC32;

Query Match 80.0%; Score 28; DB 1; Length 943;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
 |:||||  
 Db 831 KFLKFI 836

Search completed: September 7, 1999, 23:50:09  
Job time: 1948 sec

---

GenCore version 4.5

Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:36 ; Search time 148.39 Seconds  
(without alignments)  
2.903 Million cell updates/sec

Title: US-09-124-280A-7

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPREMBL10.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_prodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	91.4	140	2	032317
2	32	91.4	642	5	018203
3	31	88.6	395	1	058134
4	31	88.6	350	5	019556
5	30	85.7	260	1	027551
6	30	85.7	288	2	051841
7	30	85.7	141	2	067910
8	30	85.7	53	2	068585
9	30	85.7	493	2	067365
10	30	85.7	193	2	068417
11	30	85.7	141	2	005963
12	30	85.7	309	3	099188
13	30	85.7	1462	3	042874
14	30	85.7	640	3	093831
15	30	85.7	342	4	099633
16	30	85.7	127	5	018223
17	30	85.7	1365	5	077308
18	30	85.7	1817	5	096253
19	30	85.7	951	10	043106
20	30	85.7	953	10	P93265
21	30	85.7	1531	10	004565
22	30	85.7	956	10	042932
23	30	85.7	963	10	Q43131
24	30	85.7	956	10	Q43182
25	30	85.7	949	10	Q43243
26	30	85.7	956	10	Q43001
27	30	85.7	957	10	Q43002
28	29	82.9	312	1	058122
29	29	82.9	1279	2	026046

## ALIGNMENTS

## RESULT 1

032317  
ID O32317 PRELIMINARY; PRT; 140 AA.  
AC O32317  
DT 01-JAN-1998 (TREMREL. 05, Created)  
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)  
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)  
DE HYPOTHETICAL 15.9 KD PROTEIN.  
OS Bacillus thuringiensis.  
OG Plasmid pG13.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97405895.  
RA HOFACK L., SEURINCK J., MAHILON J.;  
RT "Nucleotide sequence and characterization of the cryptic Bacillus thuringiensis plasmid pG13 reveal a new family of rolling circle RT replicons".  
RL J. Bacteriol. 179:5000-5008(1997).  
DR EMBL; Y11173; CAAT2050.1; -.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 140 AA; 15901 MW; 48971A3F CRC32;

Query Match 91.4%; Score 32; DB 2; Length 140;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
||:||||  
Db 4 KFKVFLK 10

## RESULT 2

018203  
ID O18203 PRELIMINARY; PRT; 642 AA.  
AC O18203  
DT 01-JAN-1998 (TREMREL. 05, Created)  
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE Y48E1B.13 PROTEIN.  
GN Y48E1B.13  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MCMURRAY A.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 293393; CAB07698.1; -  
 DR PFAM: PF00655; ICE\_p10; 1.  
 DR PFAM: PF00656; ICE\_p20; 1.  
 SQ SEQUENCE 642 AA; 73590 MW; B3351DE0 CRC32;

Query Match 91.4%; Score 32; DB 5; Length 642;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

DB 157 RFLKFLK 163

RESULT 3

OS8134 PRELIMINARY; PRT; 395 AA.  
 ID O58134  
 AC O58134;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE 395AA LONG HYPOTHETICAL PROTEIN.  
 GN PH0397  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE: 98344137.  
 RA KAWABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
 RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT "Complete sequence and gene organization of the genome of a  
 RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 RL EMBL: AP000002; BAA29472.1; -  
 SQ SEQUENCE 395 AA; 46026 MW; D2DC6575 CRC32;

Query Match 88.68; Score 31; DB 1; Length 395;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

DB 237 KFLKFLK 243

RESULT 4

Q19556 PRELIMINARY; PRT; 350 AA.  
 ID Q19556  
 AC Q19556;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)

DE F18E2.4 PROTEIN.  
 GN F18E2.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LIGHTNING J.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 275537; CAA99837.1; -  
 SQ SEQUENCE 350 AA; 39851 MW; 9AA64F5B CRC32;

Query Match 88.6%; Score 31; DB 5; Length 350;  
 Best Local Similarity 85.7%; Pred. No. 92;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

DB 320 EFLKFLK 326

RESULT 5

O27551 PRELIMINARY; PRT; 260 AA.  
 ID O27551  
 AC O27551;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE 2-HYDROXYHEPTA-2,4-DIENE-1,7-DIOXIDE ISOMERASE.  
 GN MTH1507.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE: 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTIER B., QIU D.,  
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL: AE000911; AAB85982.1; -  
 KW Isomerase.  
 SQ SEQUENCE 260 AA; 28873 MW; 536811AB CRC32;

Query Match 85.7%; Score 30; DB 1; Length 260;  
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 III:II:  
 Db 5 KFLRFMK 11

RESULT 6 PRELIMINARY; PRT; 288 AA.  
 Q51841  
 ID Q51841  
 AC Q51841  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)  
 DE GPDXX, PUT, AND YHSG-PG GENES.  
 GN PUT.  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-W83;  
 RA WALLACE A.M., ROBERTS I.S.;  
 RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; X97228; CAA65865.1; -;  
 SQ SEQUENCE 288 AA; 31805 MW; 0460290B CRC32;

Query Match 85.7%; Score 30; DB 2; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 III:II:  
 Db 38 KFLSFLK 44

RESULT 7 PRELIMINARY; PRT; 141 AA.  
 O67910  
 ID O67910  
 AC O67910  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 16.6 KD PROTEIN.  
 GN AQ-2157.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000776; AAC07885.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 141 AA; 16618 MW; E880929A CRC32;

Query Match 85.7%; Score 30; DB 2; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 6  
 IIIIIII

Db 38 KFLKFLK 43

RESULT 8 PRELIMINARY; PRT; 53 AA.  
 O68585  
 ID O68585  
 AC O68585  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 6.2 KD PROTEIN.  
 OS Streptococcus mutans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-JH1005;  
 RX MEDLINE; 98261521.  
 RA HILLMAN J.D., NOVAK J., SAGURA E., GUTIERREZ J.A., BROOKS T.A.,  
 RA CROWLEY P.J., HESS M., AZIZI A., LEUNG K.-P., CVITKOVITCH D.,  
 RA BLEIWEIS A.S.;  
 RT "Genetic and biochemical analysis of mutacin 1140, a lantibiotic from  
 RT Streptococcus mutans";  
 RL Infect. Immun. 66:2743-2749(1998).  
 DR EMBL; AF051560; AAC18826.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 53 AA; 6228 MW; C2D60052 CRC32;

Query Match 85.7%; Score 30; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7  
 IIIIIII  
 Db 6 FLKFLK 11

RESULT 9 PRELIMINARY; PRT; 493 AA.  
 O67365  
 ID O67365  
 AC O67365  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 56.5 KD PROTEIN.  
 GN AQ-1350.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000736; AAC07337.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 493 AA; 56478 MW; E1603544 CRC32;

Query Match 85.7%; Score 30; DB 2; Length 493;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||  
Db 188 KFLKFL 193

## RESULT 10

O66417  
ID O66417 PRELIMINARY; PRT; 193 AA.  
AC O66417;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 22.6 KD PROTEIN.  
GN AA26.  
OS Aquifex aeolicus.  
OG Plasmid ecel.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98196666.  
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
RA GRAHAM D.E., OVERSEER R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
aeolicus.";  
RL Nature 392:353-358(1998).  
DR EMBL; AE000667; AAC07969.1; -.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 193 AA; 22552 MW; 935BAF96 CRC32;

Query Match 85.7%; Score 30; DB 2; Length 193;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7  
|||||  
Db 57 FLKFLK 62

## RESULT 11

O05963  
ID O05963 PRELIMINARY; PRT; 141 AA.  
AC O05963;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 16.5 KD PROTEIN.  
GN RP436.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=MADRID E;  
RX MEDLINE; 97419517.  
RA ANDERSSON J.O., ANDERSSON S.G.E.;  
RT "Genomic rearrangements during evolution of the obligate  
intracellular parasite Rickettsia prowazekii as inferred from an  
analysis of 52015 bp nucleotide sequence.";  
RL Microbiology 143:2783-2795(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=MADRID E;  
RX MEDLINE; 99039499.  
RA ANDERSSON S.G.E., ZOMORODIPOUR A., ANDERSSON J.O.,  
RA SICHERTZ-PONTEN T., ALSMARK U.C.M., PODOWSKI R.M., NAESLUND A.K.,  
RA ERIKSSON A.S., WINKLER H.H., KURLAND C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
mitochondria.";  
RL Nature 396:133-140(1998).

RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MADRID E;  
RA ANDERSSON S.G.E.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y11777; CAA72442.1; -.  
DR EMBL; AJ235271; CAA14893.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 16493 MW; 8AF2457A CRC32;

Query Match 85.7%; Score 30; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||  
Db 118 KFLKFL 123

## RESULT 12

O99188  
ID O99188 PRELIMINARY; PRT; 309 AA.  
AC O99188;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)  
DE ORF YOR107W FROM CHROMOSOME XV.  
GN YOR107W AND YOR3224W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomyces.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BENES V., ANDRADE M.A., RECHMANN S., TEODORU C., BANREVI A.,  
RA SANDER C., VALENCIA A., ANSORGE W., VOSS H.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA VOSS H., BENES V., RECHMANN S., TEODORU C., SCHWAGER C., PACES V.,  
RA ANSORGE W.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X94335; CAA64027.1; -.  
DR EMBL; Z75015; CAA99305.1; -.  
SQ SEQUENCE 309 AA; 35100 MW; 04E1C8F9 CRC32;

Query Match 85.7%; Score 30; DB 3; Length 309;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFL 7  
|||  
Db 50 KFLKFL 56

## RESULT 13

O42874  
ID O42874 PRELIMINARY; PRT; 1462 AA.  
AC O42874;  
DT 01-JAN-1999 (TREMBLrel. 09, Created)  
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
DE HYPOTHETICAL 164.1 KD PROTEIN C3G9.12 IN CHROMOSOME I.  
GN SPAC3G9.12.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.



```

RN  SEQUENCE FROM N.A.
RC  STRAIN-972;
RA  OLIVER K., HARRIS D., WOOD V., BARRELL B.G., RAJANDREAM M.A.;
RL  Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC  -!- SIMILARITY: TO YEAST STUL.
DR  EMBL: AL021046; CAA15921.1; -.
KW  Hypothetical protein.
FT  DOMAIN 226 229 POLY-SER.
   FT  DOMAIN 855 861 POLY-SER.
SQ  SEQUENCE 1462 AA; 164093 MW; FF20FS4E CRC32;

Query Match      85.7%; Score 30; DB 3; Length 1462;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 FLKFLK 7
   |||||
Db  9 FLKFLK 14

RESULT 14
O93831
ID  O93831 PRELIMINARY; PRT; 640 AA.
AC  O93831;
DT  01-MAY-1999 (Tremblrel. 10, Created)
DT  01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT  01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE  RAB GERANYLGERANYL TRANSFERASE ESCORT PROTEIN (REP).
GN  MRS6.
OS  Candida albicans (Yeast).
OC  Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC  Candidaceae; Candida.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-IFO1060;
RA  ISHII N., AOKI Y., ARISAWA M.;
RT  "Molecular cloning of Rab geranylgeranyl transferase escort protein
   (REP) homologue from Candida albicans.";
RL  Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AB021317; BAA36167.1; -.
KW  Transferase
SQ  SEQUENCE 640 AA; 72291 MW; C54E8D44 CRC32;

Query Match      85.7%; Score 30; DB 3; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KFLKFL 6
   |||||
Db  158 KFLKFL 163

RESULT 15
O99633
ID  O99633 PRELIMINARY; PRT; 342 AA.
AC  O99633;
DT  01-MAY-1997 (Tremblrel. 03, Created)
DT  01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT  01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE  HPRP18.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE; 97152474.
RA  HOROWITZ D.S., KRAINER A.R.;
RT  "A human protein required for the second step of pre-mRNA splicing is
   functionally related to a yeast splicing factor.";
RL  Genes Dev. 11:139-151(1997).

```

```

DR  EMBL: U51990; AAB41490.1; -.
SQ  SEQUENCE 342 AA; 39859 MW; BFD392F8 CRC32;

Query Match      85.7%; Score 30; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 KFLKFL 6
   |||||
Db  190 KFLKFL 195

```

Search completed: September 7, 1999, 20:34:37  
Job time: 19733 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:10 : Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-8  
Perfect score: 50  
Sequence: 1 KFLKFLKFLK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	32	1 R33514	Generic peptide fo
2	50	100.0	11	1 R33532	Peptide for treati
3	50	100.0	11	1 R39294	Endotoxin lipid A
4	50	100.0	10	1 R71779	Peptide neutralisi
5	50	100.0	12	1 R71791	Peptide neutralisi
6	50	100.0	11	1 W21625	Antibiotic potenti
7	50	100.0	10	1 W21396	Antibiotic potenti
8	50	100.0	12	1 W21608	Antibiotic potenti
9	45	90.0	10	1 W21629	Antibiotic potenti
10	38	76.0	10	1 R71781	Peptide neutralisi
11	38	76.0	10	1 W21598	Antibiotic potenti
12	35	70.0	7	1 R33533	Peptide for treati
13	35	70.0	7	1 R39295	Endotoxin lipid A
14	35	70.0	7	1 R71778	Peptide neutralisi
15	35	70.0	7	1 W21626	Antibiotic potenti
16	35	70.0	7	1 W21595	Antibiotic potenti
17	35	70.0	1279	1 W98761	H. pylori GPO 109
18	33	66.0	9	1 W21628	Antibiotic potenti
19	32	64.0	420	1 W19781	Tyrosyl-tRNA synth
20	32	64.0	1580	1 W57412	Homo sapiens sulph
21	31	62.0	869	1 R53732	S. cerevisiae Plc1
22	31	62.0	540	1 W20605	H. pylori cytoplas
23	31	62.0	566	1 W20945	H. pylori inner me
24	31	62.0	555	1 W84154	Human desaturase e
25	31	62.0	608	1 W84155	Human desaturase e
26	31	62.0	746	1 W84156	Human desaturase e
27	31	62.0	746	1 W85135	A desaturase enzym
28	31	62.0	555	1 W85133	A desaturase enzym
29	31	62.0	608	1 W85134	A desaturase enzym
30	31	62.0	562	1 W95512	Amino acid sequenc
31	31	62.0	615	1 W95513	Amino acid sequenc
32	31	62.0	753	1 W95514	Amino acid sequenc
33	30	60.0	6	1 R33515	Peptide for treati
34	30	60.0	1029	1 R38862	GC-A. New purified
35	30	60.0	185	1 R47115	Toxoplasma GP28.5
36	30	60.0	1077	1 R91597	C3G protein. Ras p
37	30	60.0	287	1 W06967	Cyclopropane mycol
38	30	60.0	252	1 W18315	Toxoplasma gondii
39	30	60.0	252	1 W36012	Toxoplasma gondii
40	30	60.0	468	1 W27327	Rat glial cell der
41	30	60.0	465	1 W35333	Human glial cell l
42	30	60.0	468	1 W35334	Rat glial cell lin
43	30	60.0	252	1 W32084	Non-glycosylated T

## ALIGNMENTS

### RESULT 1

R33514 ID R33514 standard; peptide; 32 AA.  
AC R33514;  
DT 07-JUL-1993 (first entry)  
DE Generic peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 2; Page 32; 39pp; English.  
CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X where n = 1-10, the peptide is useful for treating or preventing septic shock, mixing with polymyxin B to reduce its toxicity; removing endotoxins from blood, sera or other fluids (in vivo or in vitro); controlling release of cytokines induced by endotoxins; as diagnostic reagents to detect and quantify toxins in blood or sera; preparing non-toxic antigenic complexes of lipid A or lipopolysaccharide (LPS); and for treating pertussis, bacterial meningitis and HIV-related infections. The usual dose is 10-100 ug/kg/day, given parenterally. It binds to the same sites as polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It has no antibiotic activity; does not lyse erythrocytes; has no toxicity in mice when injected at 50mg/kg and is relatively unstable against proteases.  
SQ Sequence 32 AA;

Query Match 100.0%; Score 50; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLKFLK 10  
      |||||  
Db 2 KFLKFLKFLK 11

### RESULT 2

R33532 ID R33532 standard; peptide; 11 AA.  
AC R33532;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated  
 PS Claim 10; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood;  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
 Db 2 KFLKFLKFLK 11

RESULT 3  
 R3294  
 ID R3294 standard; peptide; 11 AA.  
 AC R3294;  
 DE Endotoxin (first entry)  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW Polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 50; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
 Db 1 KFLKFLKFLK 11

Db 2 KFLKFLKFLK 11

RESULT 4  
 R7179  
 ID R7179 standard; peptide; 10 AA.  
 AC R7179;  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
 Db 1 KFLKFLKFLK 10

RESULT 5  
 R71791  
 ID R71791 standard; peptide; 12 AA.  
 AC R71791;  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PH Key Location/Qualifiers  
 FT disulfide\_bond 1..12  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 22; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.

CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock CC and also in vitro to detoxify vaccines, drug solutions, injectable CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new CC peptides.  
CC Sequence 12 AA;

Query Match 100.0%; Score 50; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | |  
DB 2 KFLKFLKFLK 11

RESULT 6

W21625 ID W21625 standard; peptide; 11 AA.

AC W21625;  
DE 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #37.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
PI WPI: 97-034095/03.

DR Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 41; Page 28; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
CC Sequence 11 AA;

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | |  
DB 2 KFLKFLKFLK 11

RESULT 7

W21596 ID W21596 standard; peptide; 10 AA.

AC W21596;  
DE 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #8.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;

DR WPI: 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 12; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
CC Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | |  
DB 1 KFLKFLKFLK 10

RESULT 8

W21608 ID W21608 standard; peptide; 12 AA.

AC W21608;  
DE 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #20.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..12  
FN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
PI WPI: 97-034095/03.

DR Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 24; Page 26; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
CC Sequence 12 AA;

Query Match 100.0%; Score 50; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | |  
DB 2 KFLKFLKFLK 11

RESULT 9

W21629 ID W21629 standard; peptide; 10 AA.

AC W21629;  
DE 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #41.

KW Potentiate: antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.

PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;  
 DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 25; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 90.0%; Score 45; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9

DB 2 KFLKFLKFL 10

RESULT 10

ID R71781 standard; peptide; 10 AA.  
 AC R71781.

DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.

KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.

PN WO9503327-A.  
 PD 02-FEB-1995.

PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.

PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;

DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising

PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids

PS Claim 12; Page 21; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.

CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.

CC The present sequence is a specifically claimed example of the new  
 CC peptides.

SQ Sequence 10 AA;

Query Match

Best Local Similarity 76.0%; Score 38; DB 1; Length 10;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10

DB 1 KFFKFFKFFK 10

RESULT 11

W21598

ID W21598 standard; peptide; 10 AA.

AC W21598;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #10.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.

OS Synthetic.

PN WO9638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required

PS Claim 14; Page 25; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match

Best Local Similarity 76.0%; Score 38; DB 1; Length 10;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10

DB 1 KFFKFFKFFK 10

RESULT 12

R33533

ID R33533 standard; peptide; 7 AA.

AC R33533;

DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.

OS Synthetic.

PN ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M;

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of amino acid(s) and analogs

PS Claim 11; Page 32; 39pp; English.

CC This peptide is a specific example of a generic peptide of

CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 70.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;  
 QY 1 KFLKFLK 7  
 |||||  
 Db 1 KFLKFLK 7

RESULT 13  
 R39295  
 ID R39295 standard; peptide; 7 AA.  
 AC R39295;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 11; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins; and to detect or quantify endotoxins in blood  
 CC products. 7 AA;  
 SQ Sequence 7 AA;

Query Match 70.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;  
 QY 1 KFLKFLK 7  
 |||||  
 Db 1 KFLKFLK 7

RESULT 14  
 R71778  
 ID R71778 standard; peptide; 7 AA.  
 AC R71778;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.

OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 9; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 7 AA;  
 SQ Sequence 7 AA;

Query Match 70.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;  
 QY 1 KFLKFLK 7  
 |||||  
 Db 1 KFLKFLK 7

RESULT 15  
 W21626  
 ID W21626 standard; peptide; 7 AA.  
 AC W21626;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #38.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 03-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 42; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 70.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;  
 QY 1 KFLKFLK 7  
 |||||

Db 1 KFLKFLK 7

Search completed: September 7, 1999, 20:37:11  
Job time: 18460 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:33 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-8  
Perfect score: 50  
Sequence: 1 KFLKFLKFLK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	11	1 US-07-819-893-8	Sequence 8, Appl
2	50	100.0	11	1 US-07-819-893-8	Sequence 8, Appl
3	50	100.0	11	1 US-08-280-397-8	Sequence 8, Appl
4	50	100.0	10	1 US-08-097-8305-8	Sequence 8, Appl
5	50	100.0	12	1 US-08-097-8305-20	Sequence 20, Appl
6	50	100.0	10	2 US-08-456-112B-8	Sequence 8, Appl
7	50	100.0	12	2 US-08-456-112B-20	Sequence 20, Appl
8	50	100.0	11	2 US-08-456-112B-37	Sequence 37, Appl
9	45	90.0	10	2 US-08-456-112B-41	Sequence 41, Appl
10	38	76.0	10	1 US-08-097-8305-10	Sequence 10, Appl
11	38	76.0	10	2 US-08-456-112B-10	Sequence 10, Appl
12	35	70.0	7	1 US-07-819-893-9	Sequence 9, Appl
13	35	70.0	7	1 US-08-280-397-9	Sequence 9, Appl
14	35	70.0	7	1 US-08-097-8305-7	Sequence 7, Appl
15	35	70.0	7	2 US-08-456-112B-7	Sequence 7, Appl
16	35	70.0	7	2 US-08-456-112B-38	Sequence 38, Appl
17	33	66.0	9	2 US-08-456-112B-40	Sequence 40, Appl
18	30	60.0	287	1 US-08-457-245-3	Sequence 3, Appl
19	30	60.0	185	2 US-08-338-543-2	Sequence 2, Appl
20	30	60.0	576	2 US-08-576-279-58	Sequence 58, Appl
21	29	58.0	17	1 US-07-725-331-49	Sequence 49, Appl
22	29	58.0	17	1 US-07-725-331-53	Sequence 53, Appl
23	29	58.0	1141	1 US-08-363-300-2	Sequence 2, Appl
24	29	58.0	637	2 US-08-703-947-2	Sequence 2, Appl
25	29	58.0	238	2 US-08-380-403A-7	Sequence 7, Appl
26	29	58.0	17	3 PCT-US91-05047-49	Sequence 49, Appl
27	29	58.0	17	3 PCT-US91-05047-53	Sequence 53, Appl
28	28	56.0	22	1 US-07-965-663A-19	Sequence 19, Appl
29	28	56.0	887	1 US-08-215-709-1	Sequence 1, Appl
30	28	56.0	32	1 US-08-118-270-188	Sequence 188, App
31	28	56.0	20	1 US-08-311-611A-101	Sequence 101, App
32	28	56.0	20	1 US-08-372-783-101	Sequence 101, App
33	28	56.0	20	1 US-08-372-105-101	Sequence 101, App
34	28	56.0	20	1 US-08-306-473A-101	Sequence 101, App
35	28	56.0	116	1 US-08-702-344-21	Sequence 21, Appl
36	28	56.0	3056	2 US-08-508-836A-8	Sequence 8, Appl
37	28	56.0	20	2 US-08-621-803-148	Sequence 148, App
38	28	56.0	566	2 US-08-666-367B-5	Sequence 5, Appl
39	28	56.0	355	2 US-08-666-367B-6	Sequence 6, Appl

40 28 56.0 20 2 US-08-485-445A-101 Sequence 101, App  
41 28 56.0 3056 2 US-08-629-001A-3 Sequence 3, Appl  
42 28 56.0 1581 2 US-08-404-531B-6 Sequence 6, Appl  
43 28 56.0 1582 2 US-08-404-531B-9 Sequence 9, Appl  
44 28 56.0 1498 2 US-08-404-531B-28 Sequence 28, Appl  
45 28 56.0 1498 2 US-08-404-531B-29 Sequence 29, Appl

ALIGNMENTS

RESULT 1  
US-08-049-871-8  
; Sequence 8, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-8  
  
Query Match 100.0% Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.004; 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KFLKFLKFLK 10  
Db 2 KFLKFLKFLK 11  
  
RESULT 2  
US-07-819-893-8  
; Sequence 8, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
US-07-819-893-8

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
DB 2 KFLKFLKFLK 11

RESULT 3  
US-08-280-397-8  
;; Sequence 8, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-280-397-8

Query Match 100.0%; Score 50; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
DB 2 KFLKFLKFLK 11

RESULT 4  
US-08-097-830E-8  
;; Sequence 8, Application US/08097830E  
;; Patent No. 5652211  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Peptides For Neutralizing The  
;; TITLE OF INVENTION: Toxicity of Lipid A  
;; NUMBER OF SEQUENCES: 35  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/097,830E  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
US-08-097-830E-8

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0036;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | | | | |  
Db 1 KFLKFLKFLK 10

RESULT 5  
US-08-097-830E-20  
; Sequence 20, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys

Query Match 100.0%; Score 50; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | | | | |  
Db 2 KFLKFLKFLK 11

RESULT 6  
US-08-456-112B-8  
; Sequence 8, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York

; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-8

Query Match 100.0%; Score 50; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0036;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | | | | |  
Db 1 KFLKFLKFLK 10

RESULT 7  
US-08-456-112B-20  
; Sequence 20, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-20

Query Match 100.0%; Score 50; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
Db 2 KFLKFLKFLK 11

## RESULT 8

US-08-456-112B-37  
Sequence 37, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-37

Query Match 100.0%; Score 50; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.004;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
Db 2 KFLKFLKFLK 11

## RESULT 9

US-08-456-112B-41  
Sequence 41, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo

TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-41

Query Match 90.0%; Score 45; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
Db 2 KFLKFLKFL 10

## RESULT 10

US-08-097-830E-10  
Sequence 10, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-097-830E-10

Query Match 76.0%; Score 38; DB 1; Length 10;
Best Local Similarity 70.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10
Db 1 KFFKFFKFFK 10

RESULT 11
US-08-456-112B-10
; Sequence 10, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-10

Query Match 76.0%; Score 38; DB 2; Length 10;
Best Local Similarity 70.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10
Db 1 KFFKFFKFFK 10

RESULT 12
US-07-819-893-9
; Sequence 9, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-819-893-9

Query Match 70.0%; Score 35; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7
Db 1 KFLKFLK 7

RESULT 13
US-08-280-397-9
; Sequence 9, Application US/08280397
; Patent No. 5589459
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-280-397-9

Query Match 70.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 14  
US-08-097-830E-7  
Sequence 7, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porto, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPOID A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-7

Query Match 70.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 15  
US-08-456-112B-7  
Sequence 7, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porto, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-7

Query Match 70.0%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

Search completed: September 7, 1999, 22:38:34  
Job time: 7918 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:22 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280a-8

Perfect score: 50

Sequence: 1 KFLKFLKFLK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	76.0	394	2 F69888	alanine racemase h
2	36	72.0	956	2 A45506	H+-transporting AT
3	36	72.0	956	2 A43637	H+-transporting AT
4	36	72.0	956	2 S0751	H+-transporting AT
5	36	72.0	951	2 S52728	H+-transporting AT
6	36	72.0	949	2 T02083	H+-transporting AT
7	35	70.0	1279	2 E64709	type IIS restricti
8	34	68.0	260	2 A69068	2-hydroxyhepta-2,4
9	34	68.0	201	2 G70130	conserved hypotet
10	34	68.0	583	2 S43139	phytoene dehydroge
11	34	68.0	578	2 S55102	PS2 protein - yea
12	34	68.0	337	2 S53846	ribosomal protein
13	34	68.0	564	2 E70394	ABC transporter (h
14	34	68.0	132	2 C71095	hypothetical prote
15	33	66.0	1858	1 A44214	genome polyprotein
16	33	66.0	192	2 H70138	cell division prot
17	33	66.0	2291	2 S11238	polymerase - Berne
18	33	66.0	535	2 A64697	conserved hypotet
19	33	66.0	286	2 B70614	probable mmaal pro
20	33	66.0	140	2 H64411	hypothetical prote
21	33	66.0	4572	2 S7908	hypothetical polyp
22	33	66.0	141	2 B70485	hypothetical prote
23	33	66.0	644	2 B70420	NADH dehydrogenase
24	32.5	65.0	585	2 S11783	bla regulator prot
25	32	64.0	1866	1 GWNE2C	genome polyprotein
26	32	64.0	236	2 A56509	peroxisomal membra
27	32	64.0	936	2 B64567	cytochrome c bioge
28	32	64.0	936	2 H71862	probable cytochrom
29	32	64.0	119	2 S15218	vsdf protein - Sal
30	32	64.0	121	2 C71696	hypothetical prote
31	32	64.0	346	2 I64237	hypothetical prote
32	32	64.0	1345	2 A64430	DNA-directed RNA p
33	32	64.0	551	2 A57189	secY protein homol
34	32	64.0	206	2 S70004	hypothetical prote
35	32	64.0	553	2 T01432	secY protein homol
36	32	64.0	343	2 T01416	secY protein homol
37	32	64.0	943	2 S59317	DI22 protein - yea
38	32	64.0	110	2 S52058	hypothetical prote
39	32	64.0	342	2 S53056	probable membrane

#### ALIGNMENTS

##### RESULT 1

F69888

alanine racemase homolog yncD - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998

C:Accession: F69888

R:Kunst, S.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MUID:98044033

A:Accession: F69888

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-394 <KUN>

A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PID:el183423; PID:g2634148

A:Experimental source: strain 168

C:Genetics:

A:Gene: yncD

C:Superfamily: alanine racemase

Query Match 76.0%; Score 38; DB 2; Length 394;

Best Local Similarity 70.0%; Pred. No. 8.5;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10

DB 187 KFISFLSLK 196

##### RESULT 2

A45506

H+-transporting ATPase (EC 3.6.1.35), LHA1 - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998

C:Accession: A45506

R:Ewing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.

Plant Physiol. 94, 1874-1881, 1990

A:Title: Molecular cloning of tomato plasma membrane H+-ATPase.

A:Reference number: A45506

A:Accession: A45506

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-956 <EWI>

A:Cross-references: GB:M60166; NID:gl70463; PID:gl70464

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma

C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein

F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>

F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:427/Binding site: ATP (Lys) #status predicted

```

Query Match      72.0%; Score 36; DB 2; Length 956;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9
    ||||| |
Db 61 KFLKFLGFM 69

RESULT 3
A43637
H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco
C:Species: Nicotiana glauca (curled-leaved tobacco)
C>Date: 11-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 17-Mar-1999
C:Accession: A43637
R:Boutry, M.; Michellet, B.; Goffeau, A.
Biochem. Biophys. Res. Commun. 162, 567-574, 1989
A:Title: Molecular cloning of a family of plant genes encoding a protein homologous to P
A:Reference number: A43637; MUID:89334860
A:Accession: A43637
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-956 <BOU>
A:Cross-references: GB:M27888; NID:gl70205; PID:gl70206
A:Note: the authors translated the codon CGG for residue 131 as Trp
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:427/Binding site: ATP (Lys) #status predicted

Query Match      72.0%; Score 36; DB 2; Length 956;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9
    ||||| |
Db 61 KFLKFLGFM 69

RESULT 4
S50751
H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato
C:Species: Solanum tuberosum (potato)
C>Date: 14-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 24-Apr-1998
C:Accession: S50751; S39313
R:Harms, K.; Woehner, R.V.; Schulz, B.; Frommer, W.B.
Plant Mol. Biol. 26, 979-988, 1994
A:Title: Isolation and characterization of P-type H(+)-ATPase genes from potato.
A:Reference number: S50751
A:Accession: S50751
A:Molecule type: mRNA
A:Residues: 1-956 <HAR>
A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003
A:Note: the nucleotide is not given
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:427/Binding site: ATP (Lys) #status predicted

Query Match      72.0%; Score 36; DB 2; Length 956;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9
    ||||| |
Db 61 KFLKFLGFM 69

```

```

RESULT 5
S52728
H+-transporting ATPase (EC 3.6.1.35) - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 24-Apr-1998
C:Accession: S52728
R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.
Submitted to the EMBL Data Library, March 1995
A:Reference number: S52728
A:Accession: S52728
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-951 <CAM>
A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:485-655/Domain: ATPase nucleotide-binding domain homology <ATN>
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:425/Binding site: ATP (Lys) #status predicted

Query Match      72.0%; Score 36; DB 2; Length 951;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9
    ||||| |
Db 59 KFLKFLGFM 67

RESULT 6
T02083
H+-transporting ATPase (EC 3.6.1.35) Mhal - maize
C:Species: Zea mays (maize)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 26-Feb-1999
C:Accession: T02083
R:Jin, Y.; Bennetzen, J.L.
Plant Cell 6, 1177-1186, 1994
A:Title: Integration and nonrandom mutation of a plasma membrane proton ATPase gene f
A:Reference number: Z14555
A:Accession: T02083
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-949 <JIN>
A:Cross-references: EMBL:U09989; NID:g507770; PID:g533775
A:Experimental source: strain D3L
C:Genetics:
A:Gene: Mhal
A:Map position: 2L
A:Introns: 21/3; 61/3; 94/3; 140/3; 185/3; 245/3; 314/3; 354/3; 394/3; 435/3; 470/3;
C:Function:
A:Description: transports protons across the plasma membrane to regulate intracellular
C:Keywords: ATP; hydrogen ion transport; hydrolase; phosphoprotein

Query Match      72.0%; Score 36; DB 2; Length 949;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9
    ||||| |
Db 64 KFLKFLGFM 72

RESULT 7
E84709
type IIS restriction enzyme R and M protein - Helicobacter pylori (strain 28695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 10-Oct-1997
C:Accession: E84709
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe

```



son, J. D.; Kelley, J. M.; Cotton, M. D.; Weidman, J. M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*. A:Reference number: A64520; MUID:97394467

A:Accession: E64709

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1279 <TOM>

A:Cross-references: GB:AE000649; GB:AE000511; NID:g2314687; PID:g2314695; TIGR:HP1517

Query Match 70.0%; Score 35; DB 2; Length 1279;  
Best Local Similarity 80.0%; Pred. No. 87;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| | | | | | | |  
DB 738 KELKFKKFLK 747

RESULT 8

A:Accession: A69068

2-hydroxyhepta-2,4-diene-1,7-dioate isomerase - *Methanobacterium thermoautotrophicum* (str. Delta H)

C:Species: *Methanobacterium thermoautotrophicum*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Sep-1998

C:Accession: A69068

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: function and genome organization

A:Reference number: A69000; MUID:98037514

A:Accession: A69068

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-260 <MTH>

A:Cross-references: GB:AE000911; GB:AE000666; NID:g2622623; PID:g2622625

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1507

A:Start codon: GTG

C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 68.0%; Score 34; DB 2; Length 260;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKFLKFLK 10  
| | | | | | | |  
DB 4 LKFLKFLK 11

RESULT 9

G70130

conserved hypothetical protein BB0247 - Lyme disease spirochete

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998

C:Accession: G70130

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: G70130

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-201 <KLE>

A:Cross-references: GB:AE001135; GB:AE000783; NID:g2688144; PID:g2688151; TIGR:BB0247

A:Experimental source: strain B31

C:Superfamily: *Methanococcus jannaschii* conserved hypothetical protein MJ0226

Query Match 68.0%; Score 34; DB 2; Length 201;  
Best Local Similarity 87.5%; Pred. No. 22;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLKFLKFL 9  
| | | | | | | |  
DB 187 FLKFKFL 194

RESULT 10

S43139

phytoene dehydrogenase (EC 1.3.-.-) - *Phycomyces blakesleeanus*

N:Alternate names: phytoene desaturase

C:Species: *Phycomyces blakesleeanus*

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Sep-1997

C:Accession: S43139

R:Ruiz-Hidalgo, M.J. submitted to the EMBL Data Library, March 1994

A:Reference number: S43139

A:Accession: S43139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-583 <RUI>

A:Cross-references: EMBL:X78434; NID:g468556; PID:g468557

C:Genetics:

A:Introns: 197/2

C:Keywords: oxidoreductase

Query Match 68.0%; Score 34; DB 2; Length 583;  
Best Local Similarity 66.7%; Pred. No. 61;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLKFLKFLK 10  
| | | | | | | |  
DB 122 FLRFLDFMK 130

RESULT 11

S55102

PSP2 protein - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein YN571.01; protein YML017w

C:Species: *Saccharomyces cerevisiae*

C:Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 06-Feb-1998

C:Accession: S55102; S59358

R:Gentiles, S.; Bowman, S. submitted to the EMBL Data Library, June 1995

A:Reference number: S55102

A:Accession: S55102

A:Molecule type: DNA

A:Residues: 1-578 <GEN>

A:Cross-references: EMBL:Z49810; NID:g854472; PID:g854473; MIPS:YML017w

A:Experimental source: strain AB972

R:Formosa, T.D.; Nittis, T. submitted to the EMBL Data Library, August 1995

A:Description: High copy suppressors of the temperature sensitivity of DNA polymerase

A:Reference number: S59358

A:Accession: S59358

A:Molecule type: DNA

A:Residues: 1-578 <FOR>

A:Cross-references: EMBL:U33116; NID:g995752; PID:g995753

C:Genetics:

A:Gene: SGD:PSP2

A:Cross-references: SGD:S0004479; MIPS:YML017w

A:Map position: 13L

Query Match 68.0%; Score 34; DB 2; Length 578;  
Best Local Similarity 75.0%; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKF 8  
|| ||||  
Db 117 KFTKFKF 124

## RESULT 12

S53846

ribosomal protein S7 - Acanthamoeba castellanii mitochondrion (SGC6)

C:Species: mitochondrion Acanthamoeba castellanii

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Feb-1999

C:Accession: S53846

R:Burger, G.; Plante, I.; Lonergan, K.M.; Gray, M.W.

J. Mol. Biol. 245, 522-537, 1995

A:Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: comp

A:Reference number: S53825; MUID:95147275

A:Accession: S53846

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-337 &lt;BUR&gt;

A:Cross-references: GB:U12386; NID:g562028; PID:g562050

A:Experimental source: strain Neff; ATCC 30010

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC6

C:Keywords: mitochondrion

Query Match 68.0%; Score 34; DB 2; Length 337;  
Best Local Similarity 55.6%; Pred. No. 36;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLKFLK 10  
::|||  
Db 326 YIRFLRFLK 334

## RESULT 13

E70394

ABC transporter (hlyB subfamily) - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 18-Sep-1998

C:Accession: E70394

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V. Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666

A:Accession: E70394

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-564 &lt;AOF&gt;

A:Cross-references: GB:AE000723; NID:g2983569; PID:g2983577; GB:AE000657

A:Experimental source: strain VF5

C:Genetics:

A:Gene: abcT5

C:Superfamily: ATP-binding cassette homology

F:349-543/Domain: ATP-binding cassette homology &lt;ABC&gt;

F:366-373/Region: nucleotide-binding motif A (P-loop)

Query Match 68.0%; Score 34; DB 2; Length 564;  
Best Local Similarity 75.0%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLKF 8  
:|||||  
Db 310 EFLKFLRF 317

## RESULT 14

C71095

hypothetical protein PH1024 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998

C:Accession: C71095

R:Kawarabayashi, Y.; Sawada, M.; Horioka, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophili

A:Reference number: A71000; MUID:98344137

A:Accession: C71095

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-132 &lt;RAW&gt;

A:Cross-references: GB:AP000004; NID:g3236131; PID:d1031064; PID:g3257438

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBa

C:Genetics:

A:Gene: PH1024

Query Match 68.0%; Score 34; DB 2; Length 132;  
Best Local Similarity 70.0%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| ||| |||  
Db 121 KMLKFKFLR 130

## RESULT 15

A44214

genome polyprotein 1 - cowpea severe mosaic virus

N:Contains: 24K protein; 32K protein; 58K membrane-binding protein; RNA-directed RNA

C:Species: cowpea severe mosaic virus

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 31-Dec-1996

C:Accession: A44214

R:Chen, X.; Bruening, G.

Virology 191, 607-618, 1992

A:Title: Cloned DNA copies of cowpea severe mosaic virus genomic RNAs: infectious tra

A:Reference number: A44214; MUID:93079863

A:Accession: A44214

A:Molecule type: genomic RNA

A:Residues: 1-1858 &lt;CHE&gt;

A:Cross-references: GB:M83830

C:Genetics:

A:Map position: segment 1

C:Superfamily: cowpea aphid-borne mosaic virus genome polyprotein B

C:Keywords: genome-linked protein; glycoprotein; hydrolase; nucleotidyltransferase; p

F:2-313/Product: 32K proteinase cofactor #status predicted &lt;PCF&gt;

F:218-235/Region: hydrophobic

F:314-908/Product: 58K membrane-binding protein #status predicted &lt;MBP&gt;

F:891-907/Domain: transmembrane #status predicted &lt;TMN&gt;

F:909-936/Product: genome-linked protein VPg #status predicted &lt;VPG&gt;

F:937-1146/Product: 24K viral proteinase #status predicted &lt;VPT&gt;

F:1147-1858/Product: RNA-directed RNA polymerase #status predicted &lt;RRP&gt;

F:1811-1826/Region: hydrophobic

F:326.958,1744,1758/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.0%; Score 33; DB 1; Length 1858;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
|||||:  
Db 1379 KFLKVVVFI 1387

Search completed: September 7, 1999, 23:06:23

Job time: 2471 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:09 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-8

Perfect score: 50

Sequence: 1 KFLKFLKFLK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	76.0	394	1	ALR2_BACSU
2	36	72.0	956	1	PMAL_LYCES
3	34	68.0	583	1	CR21_PHYBL
4	34	68.0	578	1	PSP2_YEAST
5	34	68.0	103	1	RR6_CVACA
6	34	68.0	337	1	RT07_ACACA
7	33	66.0	140	1	FLAE_METJA
8	33	66.0	2291	1	RRPB_BEV
9	33	66.0	1858	1	VGNB_CPMV
10	32.5	65.0	585	1	BLAR_STAAT
11	32	64.0	943	1	DIP2_YEAST
12	32	64.0	235	1	PEXB_YEAST
13	32	64.0	1341	1	RPAL_METJA
14	32	64.0	1580	1	SUR_HUMAN
15	32	64.0	1866	1	VGNB_CPMV
16	32	64.0	119	1	VSPF_SALDU
17	32	64.0	346	1	Y343_MYCGE
18	32	64.0	110	1	YJ20_YEAST
19	32	64.0	342	1	YMI2_YEAST
20	32	64.0	824	1	YOT5_CAEEL
21	31	62.0	762	1	MSU2_HELPY
22	31	62.0	979	1	P115_MYCHR
23	31	62.0	869	1	PLC1_YEAST
24	31	62.0	179	1	RL15_ARCFU
25	31	62.0	2210	1	RRFO_TACV
26	31	62.0	233	1	VENV_BEV
27	31	62.0	729	1	VGNB_APMV
28	31	62.0	1659	1	VIT_ONCMY
29	31	62.0	309	1	VP62_MRDV
30	31	62.0	1327	1	YGL2_YEAST
31	31	62.0	1122	1	YGC3_YEAST
32	31	62.0	922	1	YK16_CAEEL
33	31	62.0	665	1	YL14_CAEEL
34	30.5	61.0	238	1	Y037_METJA
35	30	60.0	1061	1	ANPA_HUMAN
36	30	60.0	1057	1	ANPA_MOUSE
37	30	60.0	1057	1	ANPA_RAT
38	30	60.0	287	1	CFAL_MYCTU
39	30	60.0	500	1	CP03_RAT
40	30	60.0	1630	1	ESPL_YEAST
41	30	60.0	2747	1	FAF_DRONE
42	30	60.0	464	1	GDNR_HUMAN
43	30	60.0	468	1	GDNR_MOUSE

ALIGNMENTS

RESULT 1

ALR2\_BACSU  
ID ALR2\_BACSU STANDARD; PRT: 394 AA.  
AC P94494;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DE 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE PUTATIVE ALANINE RACEMASE (EC 5.1.1.1).  
GN YNCD.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BORCHERT S., KLEIN C., PIKSA B., HAMMELMANN M., ENTIAN K.D.;  
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC 1- CATALYTIC ACTIVITY: L-ALANINE - D-ALANINE.  
CC 1- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC 1- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: U66480; GI750128; -  
CC EMBL: X99113; E1183423; -  
CC SUBTILIST; BG12267; YNCD.  
CC PROSITE; PS00395; ALANINE\_RACEMASE; 1.  
CC PFAM; PF00842; ALA\_racemase; 1.  
KW HYPOTHETICAL PROTEIN; ISOMERASE; PYRIDOXAL PHOSPHATE.  
FT BINDING 39 39  
FT SEQUENCE 394 AA; 43648 MW; B5C38A51 CRC32;

Query Match 76.0%; Score 38; DB 1; Length 394;  
Best Local Similarity 70.0%; Pred.No. 3.9;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
||: || || ||  
DB 187 KETISFLSFLK 196

RESULT 2

PMAL\_LYCES  
ID PMAL\_LYCES STANDARD; PRT: 956 AA.  
AC P22180;  
DT 01-AUG-1991 (REL. 19, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).  
GN LHA1.  
OS LYCOPERSICON ESCULENTUM (TOMATO).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC ASTERIDAE; SOLANACEAE; SOLANACEAE; SOLANACEAE; SOLANACEAE;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. CASTELMAR; TISSUE=ROOT;  
RA EWING N.N., WIMMERS L.E., MEYER D.J., CHETELAT R.T., BENNETT A.B.;  
RT "Molecular cloning of tomato plasma membrane H<sup>+</sup>-ATPase."  
PLANT PHYSIOL. 94:1874-1881(1990).



```

DR EMBL: U33116; G955753; -
DR EMBL: U29398; G1808701; -
DR EMBL: 249810; G854473; -
DR SGD: L0002882; PSP2.
FT DOMAIN 393 402 POLY-ASN.
FT DOMAIN 438 453 POLY-ASN.
FT DOMAIN 457 460 POLY-ASN.
SQ SEQUENCE 578 AA; 64020 MW; 7DA753A1 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 578;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 8
Db 117 KFTKFKF 124

RESULT 5
RR6_CYACA STANDARD; PRT; 103 AA.
AC O19917;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S6.
GN RPS6.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OC CHLOROPLAST.
OC EUKARYOTA.
OC BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDACEAE;
OC CYANIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK-1;
RA GLOCKNER G., ROSENTHAL A., VALENTIN K.;
RT "Organisation of 46 kb of the Cyanidium caldarium RK1 plastid
genome.";
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: BINDS TOGETHER WITH S18 TO 16S RIBOSOMAL RNA
(BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U12386; G562050; -
CC PROSITE: PS00052; RIBOSOMAL_S7; FALSE_NEG.
CC PFAM: PF00177; S7; 1.
CC RIBOSOMAL PROTEIN; MITOCHONDRION.
SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 337;
Best Local Similarity 55.6%; Pred. No. 17;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLKFLK 10
Db 326 YIRFLRFLK 334

RESULT 7
FLAE_METJA
ID FLAE_METJA STANDARD; PRT; 140 AA.
AC Q58306;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PUTATIVE FLAGELLA-RELATED PROTEIN E.
GN FLAE OR MJ0896.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., RAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: STRONG, TO M.VOLTAE FLAE, ALSO TO FLAD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U67533; G1591572; -  
 DR TIGR; MJ0896; -  
 KW FLAGELLA.  
 SO SEQUENCE 140 AA; 15891 MW; 61446FEB CRC32;

Query Match 66.0%; Score 33; DB 1; Length 140;  
 Best Local Similarity 70.0%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
 | | | | | : |  
 Db 73 KLLKFLKGK 82

RESULT 8  
 RRPB\_BEV STANDARD; PRT; 2291 AA.  
 AC P18458;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (ORF1B).  
 GN POL.  
 OS BERNE VIRUS (BEV).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; NIDOVIRALES;  
 CC CORONAVIRIDAE; TOROVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ISOLATE P138/72;  
 RX MEDLINE; 90356389.  
 RA SNIJDER E.J., DEN BOON J.A., BREDEBEEK P.J., HORZINEK M.C.,  
 RA RIJNBAND R., SPAAN W.J.M.;  
 RT "The carboxyl-terminal part of the putative Berne virus polymerase is  
 RT expressed by ribosomal frameshifting and contains sequence motifs  
 RT which indicate that toro- and coronaviruses are evolutionarily  
 related.";  
 RL NUCLEIC ACIDS RES. 18:4535-4542(1990).  
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE  
 CC + RNA(N).  
 CC -1- THIS PROTEIN IS EXPRESSED BY AN EFFICIENT RIBOSOMAL FRAMESHIFTING  
 CC MECHANISM. RIBOSOMAL FRAMESHIFTING IS AN ELEGANT MECHANISM FOR  
 CC REGULATING THE SYNTHESIS OF SEVERAL PROTEINS IN A WELL BALANCED  
 CC MANNER.  
 CC -1- SIMILARITY: TO CORONAVIRUSES POLYMERASE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X52374; E23105; -  
 DR PIR; S11238; S11238.  
 KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
 FT DOMAIN 515 744  
 FT POLYMERASE.  
 FT DOMAIN 1099 1374  
 FT NP\_BIND 1101 1108  
 FT ATP (BY SIMILARITY).  
 FT SIMILAR 845 928  
 FT TO ITS HOMOLOG FROM MHV (AA 938-1022)  
 FT AND IBV (AA 929-1013).  
 FT SIMILAR 1921 2209  
 FT TO ITS HOMOLOG FROM MHV (AA 2322-2641)  
 FT AND IBV (AA 2274-2598).  
 FT SEQUENCE 2291 AA; 260843 MW; C0283C2D CRC32;

Query Match 66.0%; Score 33; DB 1; Length 2291;  
 Best Local Similarity 87.5%; Pred. No. 1.7e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LKFLKFLK 10  
 | | | | | : |  
 Db 523 LKFSKFLK 530

RESULT 9  
 VGNB\_CPSMV STANDARD; PRT; 1858 AA.  
 AC P36312;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN B [CONTAINS: PROTEASE COFACTOR; MEMBRANE BINDING  
 DE PROTEIN; VPG; PROTEASE (EC 3.4.22.-); RNA POLYMERASE (EC 2.7.7.48)].  
 OS COWPEA SEVERE MOSAIC VIRUS (STRAIN DG) (CPSMV).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; COMOVIRIDAE;  
 CC COMOVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93079863.  
 RA CHEN X., BRUENING G.;  
 RT "Cloned DNA copies of cowpea severe mosaic virus genomic RNAs:  
 RT Infectious transcripts and complete nucleotide sequence of RNA 1.";  
 RL VIROLOGY 191:607-618(1992).  
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M83830; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A44214; A44214.  
 DR PFAM; PF00860; RNA\_dep\_RNA\_pol; 1.  
 DR PFAM; PF00910; RNA\_helicase; 1.  
 KW POLYPROTEIN; TRANSMEMBRANE; HYDROLASE; PROTEASE; TRANSFERASE;  
 CC RNA-DIRECTED RNA POLYMERASE; ATP-BINDING.  
 FT CHAIN 1 313  
 FT PROTEASE COFACTOR (POTENTIAL).  
 FT CHAIN 314 908  
 FT MEMBRANE-BINDING PROTEIN (POTENTIAL).  
 FT CHAIN 909 936  
 FT VPG PROTEIN (POTENTIAL).  
 FT CHAIN 937 1146  
 FT PROTEASE (POTENTIAL).  
 FT CHAIN 1147 1858  
 FT RNA-DIRECTED RNA POLYMERASE (POTENTIAL).  
 FT TRANSMEM 891 907  
 FT POTENTIAL.  
 FT NP\_BIND 482 489  
 FT ATP (POTENTIAL).  
 FT ACT\_SITE 976 976  
 FT THIOL PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1013 1013  
 FT THIOL PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1104 1104  
 FT THIOL PROTEASE (POTENTIAL).  
 SQ SEQUENCE 1858 AA; 208861 MW; 38B826AB CRC32;

Query Match 66.0%; Score 33; DB 1; Length 1858;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 | | | | | : |  
 Db 1379 KFLKFWVF 1387

RESULT 10  
 BLAR\_STAAU STANDARD; PRT; 585 AA.  
 ID BLAR\_STAAU  
 AC P18357;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE REGULATORY PROTEIN BLAR1.

GN BLAR1.  
OS STAPHYLOCOCCUS AUREUS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC STAPHYLOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NCIC 9789; TRANSPOSON-TN552;  
RX MEDLINE: 91014696.  
RA ROWLAND S.-J., DYKE K.G.H.;  
RT "TN552, a novel transposable element from Staphylococcus aureus.";  
RL MOL. MICROBIOL. 4:961-975(1990).  
CC -!- FUNCTION: BLAR1 IS A POTENTIAL PENICILLIN-BINDING PROTEIN REQUIRED  
CC FOR INDUCTION OF BETA-LACTAMASE.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X52734; G46759; -  
CC PIR: S11783; S11783.  
CC DR PIR: S27371; S27371.  
CC DR PFAM: PF00144; beta-lactamase; 1.  
CC KW TRANSPOSABLE ELEMENT; TRANSMEMBRANE.  
CC FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 4 22 POTENTIAL.  
CC FT DOMAIN 23 37 EXTRACELLULAR (POTENTIAL).  
CC FT TRANSMEM 38 58 POTENTIAL.  
CC FT DOMAIN 59 107 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 108 128 POTENTIAL.  
CC FT DOMAIN 129 213 EXTRACELLULAR (POTENTIAL).  
CC FT TRANSMEM 214 233 POTENTIAL.  
CC FT DOMAIN 234 308 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 309 330 POTENTIAL.  
CC FT DOMAIN 331 585 EXTRACELLULAR (POTENTIAL).  
CC FT DOMAIN 331 585 EXTRACELLULAR (POTENTIAL).  
CC FT ACT\_SITE 389 PENICILLIN-BINDING (BY SIMILARITY).  
CC SEQUENCE 585 AA; 69245 MW; 138663BD CRC32;  
SQ

Query Match 65.0%; Score 32.5; DB 1; Length 585;  
Best Local Similarity 61.5%; Pred. No. 54;  
Matches 8; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 1 KFLK---FLKFLK 10  
| | | | : | | | |  
Db 120 KFLKALLYLKYLK 132

RESULT 11  
DIP2\_YEAST  
ID DIP2\_YEAST STANDARD; PRT; 943 AA.  
AC Q12220; Q05386;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DOM34 INTERACTING PROTEIN 2.  
GN DIP2 OR YLR129W OR L9233.1 OR L3116.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,

RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
RA WILSON R., WATERSTON R.;  
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA DELIUS H.;  
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 1-458 FROM N.A.  
RC STRAIN-FY23 / RD005;  
RA VERHASSELT P., VOET M., VOLCKAERT G.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
CC -!- SIMILARITY: TO S.POMBE SPBC3D5.12.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U53877; G1256869; -  
CC DR EMBL: U53881; G2580465; -  
CC DR EMBL: X91258; G995691; -  
CC DR EMBL: Z73301; E245568; -  
CC DR EMBL: Z73302; E322129; -  
CC DR EMBL: X89514; E198829; -  
CC SGD: L0003088; DIP2.  
CC DR PROSITE: PS00678; WD\_REPEATS; 4.  
CC PFAM: PF00400; G-beta; 8.  
CC KW REPEAT; WD\_REPEAT.  
CC FT REPEAT 77 107 WD1.  
CC FT REPEAT 119 149 WD2.  
CC FT REPEAT 161 190 WD3.  
CC FT REPEAT 202 230 WD4.  
CC FT REPEAT 389 418 WD5.  
CC FT REPEAT 428 458 WD6.  
CC FT REPEAT 471 501 WD7.  
CC FT REPEAT 571 601 WD8.  
CC FT REPEAT 613 643 WD9.  
CC FT REPEAT 655 685 WD10.  
CC SEQUENCE 943 AA; 106342 MW; 7B348EA8 CRC32;  
SQ

Query Match 64.0%; Score 32; DB 1; Length 943;  
Best Local Similarity 85.7%; Pred. No. 11e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKFLKFL 9  
| | | | | | | | | |  
Db 830 LKFLKFL 836

RESULT 12  
PEXB\_YEAST  
ID PEXB\_YEAST STANDARD; PRT; 235 AA.  
AC Q12462;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PEROXISOMAL MEMBRANE PROTEIN PMP27 (PEROXIN-11).  
GN PEX11 OR PMP27 OR PMP24 OR YOL147C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.  
RC STRAIN-S288C;  
RA MEDLINE: 95164555.  
RA ERDMANN R., BLOBEL G.;

RT "Giant peroxisomes in oleic acid-induced *Saccharomyces cerevisiae*  
 RL lacking the peroxisomal membrane protein Pmp27p.";  
 RJ J. CELL BIOL. 128:509-523(1995).  
 [2]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-29 AND 191-206.  
 RP STRAIN-GRF88;  
 RX MEDLINE; 95238534.  
 RA MARSHALL P.A., KRIMKEVICH Y.I., LARK R.H., DYER J.M., VEENHUIS M.,  
 RJ GOODMAN J.M.;  
 RT "Pmp27 promotes peroxisomal proliferation.";  
 RL J. CELL BIOL. 129:345-355(1995).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-S288C / FY1679;  
 RX MEDLINE; 96132030.  
 RA CASAMAYOR A., ALDEA M., CASAS C., ARINO J.,  
 RJ LAFUENTE M.J., GANCEDO C., HERRERO E., GAMO F.J.,  
 RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast  
 RL chromosome XV containing seven new open reading frames.";  
 RJ YEAST 11:1281-1288(1995).  
 CC -!- FUNCTION: INVOLVED IN PEROXISOMAL PROLIFERATION. COULD PARTICIPATE  
 CC IN PEROXISOMAL ELONGATION OR FISSION. MAY BE INVOLVED IN PARCELING  
 CC OF PEROXISOMES INTO REGULAR QUANTA.  
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL MEMBRANE-ASSOCIATED.  
 CC -!- SIMILARITY: STRONG, TO C.BOIDINII PEX11.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z48239; G886949; -  
 DR EMBL; Z74889; E251925; -  
 DR EMBL; X81465; G683701; -  
 DR EMBL; Z46846; G791132; -  
 DR SGD; L0002650; PMP27.  
 KW PEROXISOME; MEMBRANE.  
 FT INIT\_MET 0  
 SQ SEQUENCE 235 AA; 26744 MW; 228BDA38 CRC32;

Query Match 64.0%; Score 32; DB 1; Length 235;  
 Best Local Similarity 85.7%; Pred. No. 27;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KFLKFLK 7  
 |||:||||  
 DB 59 KFLRFLK 65

RESULT 13  
 RPAL\_METJA STANDARD; PRT; 1341 AA.  
 AC Q58445;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DNA-DIRECTED RNA POLYMERASE SUBUNIT A' (EC 2.7.7.6).  
 GN RPOA1 OR MJ1042.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EUKARYOTAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAJAVE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,

RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RJ *jannaschii*.";  
 RL SCIENCE 273:1058-1073(1996).  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N).  
 CC -!- COFACTOR: ZINC.  
 CC -!- SIMILARITY: THE COMBINED A'+A" SUBUNITS CORRESPOND TO THE A  
 CC SUBUNIT OF EUKARYOTIC RNA POLYMERASES I, II AND III AND TO THE  
 CC EUBACTERIAL BETA' SUBUNIT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U67547; G2826359; -  
 DR TIGR; MJ1042; -  
 DR PFAM; PF00623; RNA\_pol\_A; 1.  
 KW TRANSCRIPTION; DNA-DIRECTED RNA POLYMERASE; ZINC.  
 FT ZN\_FING 62 105 BY SIMILARITY.  
 SQ SEQUENCE 1341 AA; 152780 MW; 6667A582 CRC32;  
 Query Match 64.0%; Score 32; DB 1; Length 1341;  
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFLKFLKFLK 10  
 ||:| ||:|  
 DB 857 KFTFDKFK 866  
 RESULT 14  
 SUR\_HUMAN STANDARD; PRT; 1580 AA.  
 ID Q09428;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE SULFONYLUREA RECEPTOR.  
 GN SUR.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA GONZALEZ G., AGUILAR-BRYAN L., BRYAN J.;  
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
 [2]  
 RN SEQUENCE OF 1186-1580 FROM N.A.  
 RC TISSUE-PANCREATIC ISLETS;  
 RX MEDLINE; 95232533.  
 RA THOMAS P.M., COTE G.J., WOHLK N., HADDAD B., MATHEW P.M., RABL W.,  
 RA AGUILAR-BRYAN L., GAGEL R.F., BRYAN J.;  
 RT "Mutations in the sulfonylurea receptor gene in familial persistent  
 RT hyperinsulinemic hypoglycemia of infancy.";  
 RL SCIENCE 268:426-429(1995).  
 [3]  
 RN VARIANT PHFI VAL-717.  
 RP MEDLINE; 96354544.  
 RX THOMAS P.M., WOHLK N., HUANG E., KUHNLE U., RABL W., GAGEL R.F.,  
 RA COTE G.J.;  
 RT "Inactivation of the first nucleotide-binding fold of the



DR	EMBL; L78214; G1374920; JOINED.	
DR	EMBL; L78215; G1374920; JOINED.	
DR	EMBL; L78216; G1374920; JOINED.	
DR	EMBL; L78217; G1374920; JOINED.	
DR	EMBL; L78218; G1374920; JOINED.	
DR	EMBL; L78219; G1374920; JOINED.	
DR	EMBL; L78220; G1374920; JOINED.	
DR	EMBL; L78221; G1374920; JOINED.	
DR	EMBL; L78222; G1374920; JOINED.	
DR	EMBL; L78223; G1374920; JOINED.	
DR	EMBL; L78224; G1374920; JOINED.	
DR	EMBL; L78225; G1374920; JOINED.	
DR	EMBL; L78226; G1374920; JOINED.	
DR	EMBL; L78227; G1374920; JOINED.	
DR	EMBL; L78228; G1374920; JOINED.	
DR	EMBL; L78229; G1374920; JOINED.	
DR	EMBL; L78230; G1374920; JOINED.	
DR	EMBL; L78231; G1374920; JOINED.	
DR	EMBL; L78232; G1374920; JOINED.	
DR	EMBL; L78233; G1374920; JOINED.	
DR	EMBL; L78234; G1374920; JOINED.	
DR	EMBL; L78235; G1374920; JOINED.	
DR	EMBL; L78236; G1374920; JOINED.	
DR	EMBL; L78237; G1374920; JOINED.	
DR	EMBL; L78238; G1374920; JOINED.	
DR	EMBL; L78239; G1374920; JOINED.	
DR	EMBL; L78240; G1374920; JOINED.	
DR	EMBL; L78241; G1374920; JOINED.	
DR	EMBL; L78242; G1374920; JOINED.	
DR	EMBL; L78207; G1369844; -	
DR	EMBL; L40625; G784882; -	
DR	MIM; 600509; -	
DR	MIM; 256450; -	
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.	
DR	PFAM; PF00005; ABC_tran; 2.	
DR	PFAM; PF00664; ABC_membrane; 2.	
DR	HSP; P13565; INBD.	
KW	RECEPTOR; TRANSMEMBRANE; ATP-BINDING; PHOSPHORYLATION; -	
KW	ALTERNATIVE SPLICING; DISEASE MUTATION.	
FT	INIT_MET 0 0 BY SIMILARITY.	
FT	DOMAIN 1 29 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 30 50 POTENTIAL.	
FT	DOMAIN 51 73 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 74 93 POTENTIAL.	
FT	DOMAIN 94 137 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 138 133 POTENTIAL.	
FT	DOMAIN 154 167 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 168 188 POTENTIAL.	
FT	DOMAIN 189 305 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 306 323 POTENTIAL.	
FT	DOMAIN 324 349 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 350 367 POTENTIAL.	
FT	DOMAIN 368 447 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 448 470 POTENTIAL.	
FT	DOMAIN 471 539 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 540 560 POTENTIAL.	
FT	DOMAIN 561 575 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 576 586 POTENTIAL.	
FT	DOMAIN 597 1000 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 1001 1020 POTENTIAL.	
FT	DOMAIN 1021 1062 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 1063 1076 POTENTIAL.	
FT	DOMAIN 1077 1133 CYTOPLASMIC (POTENTIAL).	
FT	TRANSEM 1154 1173 POTENTIAL.	
FT	DOMAIN 1174 1275 EXTRACELLULAR (POTENTIAL).	
FT	TRANSEM 1276 1297 POTENTIAL.	
FT	DOMAIN 1298 1580 CYTOPLASMIC (POTENTIAL).	
FT	NP_BIND 712 719 ATP (POTENTIAL).	
FT	NP_BIND 1377 1384 ATP (POTENTIAL).	
FT	CARBOHYD 9 9 POTENTIAL.	
FT	VARSPIC 739 739 S -> SS (IN 2ND FORM).	
FT	VARIANT 717 717 S -> V (IN PHET).	
SO	SEQUENCE 1580 AA: 176876 MW: 86570A2 CRC32:	

Query Match 64.0%; Score 32; DB 1; Length 1580;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 | :||:|  
 Db 149 KTIKVRFL 157

RESULT 15  
 VGNB\_CPMV STANDARD; PRT; 1866 AA.  
 AC P03600;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN B [CONTAINS: PROTEASE COFACTOR; MEMBRANE BINDING  
 DE PROTEIN; VPG; PROTEASE (EC 3.4.22.-); RNA POLYMERASE (EC 2.7.7.48)].  
 OS COMPEA MOSAIC VIRUS (CPMV).  
 OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; COMOVIRIDAE;  
 OC COMOVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LOMONOSOFF G.P., SHANKS M.;  
 RT "The nucleotide sequence of cowpea mosaic virus B RNA."  
 RL ENBO J. 2:2253-2258(1983).  
 CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -!- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X00206; G58909; -  
 DR PIR; A04211; GNWEC.  
 DR PFAM; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR PFAM; PF00910; RNA\_helicase; 1.  
 KW POLYPROTEIN; TRANSMEMBRANE; HYDROLASE; PROTEASE; TRANSFERASE;  
 KW RNA-DIRECTED RNA POLYMERASE; ATP-BINDING.  
 FT CHAIN 1 326 PROTEASE COFACTOR (POTENTIAL).  
 FT CHAIN 327 919 MEMBRANE-BINDING PROTEIN (POTENTIAL).  
 FT CHAIN 920 947 VPG PROTEIN (POTENTIAL).  
 FT CHAIN 948 1155 PROTEASE (POTENTIAL).  
 FT CHAIN 1156 1866 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).  
 FT NP\_BIND 494 501 ATP (POTENTIAL).  
 FT ACT\_SITE 987 987 THIOL PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1023 1023 THIOL PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1113 1113 THIOL PROTEASE (POTENTIAL).  
 SQ SEQUENCE 1866 AA; 209809 MW; 75077012 CRC32;

Query Match 64.0%; Score 32; DB 1; Length 1866;  
 Best Local Similarity 55.6%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 ||| |:::  
 Db 1389 KFLNVRFI 1397

Search completed: September 7, 1999, 23:50:11  
 Job time: 1950 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:37 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: us-09-124-280a-8  
Perfect score: 50  
Sequence: 1 KFLKFLKFLK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_prodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	72.0	951	10 Q43106	Q43106 phaseolus v
2	36	72.0	953	10 P93265	P93265 mesembryant
3	36	72.0	956	10 Q42932	Q42932 nicotiana p
4	36	72.0	963	10 Q43131	Q43131 vicia faba
5	36	72.0	956	10 Q43182	Q43182 solanum tub
6	36	72.0	949	10 Q43243	Q43243 zea mays (m
7	36	72.0	956	10 Q43001	Q43001 oryza sativ
8	36	72.0	957	10 Q43002	Q43002 oryza sativ
9	36	72.0	403	13 Q73742	Q73742 xenopus lae
10	35	70.0	1279	2 Q26046	Q26046 helicobacte
11	35	70.0	290	3 P78815	P78815 schizosacch
12	35	70.0	552	8 Q92YM7	Q92YM7 rhizopcephal
13	34	68.0	260	1 Q27551	Q27551 methanobact
14	34	68.0	132	1 Q58773	Q58773 pyrococcus
15	34	68.0	201	2 Q51263	Q51263 borrelia bu
16	34	68.0	584	2 Q67184	Q67184 aquifex aeo
17	34	68.0	350	5 Q19556	Q19556 caenorhabdi
18	34	68.0	873	5 Q18228	Q18228 caenorhabdi
19	33	66.0	192	2 Q51293	Q51293 borrelia bu
20	33	66.0	286	2 P94922	P94922 mycobacteri
21	33	66.0	644	2 Q67389	Q67389 aquifex aeo
22	33	66.0	141	2 Q67910	Q67910 aquifex aeo
23	33	66.0	1289	5 Q17174	Q17174 boophilus m
24	33	66.0	72	5 Q23632	Q23632 caenorhabdi
25	33	66.0	93	5 Q01587	Q01587 caenorhabdi
26	33	66.0	649	5 Q4752	Q4752 caenorhabdi
27	33	66.0	4572	10 Q40712	Q40712 oryza sativ
28	32.5	65.0	585	2 Q00419	Q00419 staphylococ
29	32	64.0	450	1 Q59355	Q59355 pyrococcus

## ALIGNMENTS

### RESULT 1

Q43106 PRELIMINARY; PRT; 951 AA.  
AC Q43106;  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
GN BHA-1.  
OS Phaseolus vulgaris (Kidney bean) (French bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
OC Phaseolus.  
RN [1]  
SEQUENCE FROM N.A.

RA CAMPOS F., PEREZ-CASTINEIRA J.R., VILLALBA J.M., SERRANO R.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X85804; CAA59799.1; -  
DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
DR MENDEL; 13288; PHAVU:494;1.  
DR PFAM; PF00122; E1-E2\_ATPase; 1.  
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
FT MOD\_RES 331 331 PHOSPHORYLATION (PROBABLE).  
SQ SEQUENCE 951 AA; 104442 MW; 3B569E8C CRC32;

Query Match 72.0%; Score 36; DB 10; Length 951;  
Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
DB 59 KFLKFLGFM 67

### RESULT 2

P93265 PRELIMINARY; PRT; 953 AA.  
AC P93265;  
DT 01-MAY-1997 (TREMELrel. 03, Created)  
DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)  
DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
GN PMA.  
OS Mesembryanthemum crystallinum (Common ice plant).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Caryophyllales; Alzooceae;  
OC Mesembryanthemum.  
RN [1]  
SEQUENCE FROM N.A.

RA MICHALOWSKI C.B., QUIGLEY-LANDREAU F., BOHNERT H.J.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U84891; AAB41898.1; -  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13177; MESCL:494; 1.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 335 335 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 953 AA; 104924 MW; B11E9AA4 CRC32;

Query Match 72.0%; Score 36; DB 10; Length 953;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 ||||| I;  
 DB 63 KFLKFLGFM 71

RESULT 3  
 Q42932 PRELIMINARY; PRT; 956 AA.  
 ID Q42932;  
 AC Q42932;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE.

OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Nicotiana.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89334860.  
 RA BOUTRY M., MICHELET B., GOFFEAU A.;  
 RT "Molecular cloning of a family of plant genes encoding a protein  
 homologous to plasma membrane H<sup>+</sup>-translocating ATPases.";  
 RL Biochem. Biophys. Res. Commun. 162:567-574 (1989).  
 DR EMBL; M27888; AAA34052.1; -  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 16027; NICPL:494.mnl6027.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105047 MW; FF2D51BE CRC32;

Query Match 72.0%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 ||||| I;  
 DB 61 KFLKFLGFM 69

RESULT 4  
 Q43131 PRELIMINARY; PRT; 963 AA.  
 ID Q43131;  
 AC Q43131;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H(+)-ATPASE.

OS Vicia faba (Broad bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Vicia.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 96012934.  
 RA NAKAJIMA N., SATI H., AONO M., KONDO N.;  
 RT "Isolation of cDNA for a plasma membrane H(+)-ATPase from guard cells  
 of Vicia faba L.";  
 RL Plant Cell Physiol. 36:919-924 (1995).  
 DR EMBL; S79323; AAB35314.1; -  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13431; VICFA:494; 3.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 338 338 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 963 AA; 105965 MW; 2C2C9793 CRC32;

Query Match 72.0%; Score 36; DB 10; Length 963;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 ||||| I;  
 DB 66 KFLKFLGFM 74

RESULT 5  
 Q43182 PRELIMINARY; PRT; 956 AA.  
 ID Q43182;  
 AC Q43182;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
 GN PHAL.

OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Solanum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LEAF;  
 RX MEDLINE; 95093037.  
 RA HARMS K., WOHNER R.V., SCHULZ B., FROMMER W.B.;  
 RT "Isolation and characterization of P-type H(+)-ATPase genes from  
 potato.";  
 RL Plant Mol. Biol. 26:979-988 (1994).  
 DR EMBL; X76536; CAA54046.1; -  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 530; SOLTU:494; 1.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105125 MW; 80AD1D45 CRC32;

Query Match 72.0%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 ||||| I;  
 DB 61 KFLKFLGFM 69

RESULT 6  
 Q43243 PRELIMINARY; PRT; 949 AA.  
 ID Q43243;  
 AC Q43243;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE.  
 GN MH1.

OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Zea.  
 [1]  
 RN PARTIAL SEQUENCE FROM N.A.  
 RC STRAIN=D3L;  
 RX MEDLINE; 95003707.  
 RA JIN Y., BENNETZEN J.L.;  
 RT "Integration and nonrandom mutation of a plasma membrane proton  
 AT ATPase gene fragment within the Bsl retroelement of maize.";  
 RL Plant Cell 6:1177-1186(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D3L;  
 RA JIN Y.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U09989; AAB60276.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 533; ZEAMa; 494; 2.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 336 336 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 949 AA; 103900 MW; DDIAAOCF CRC32;

Query Match 72.0%; Score 36; DB 10; Length 949;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 DB 64 KFLKFLGFM 72  
 ||||| |

RESULT 7  
 Q43001 ID Q43001 PRELIMINARY; PRT; 956 AA.  
 AC Q43001;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE H-ATPASE (EC 3.6.1.3) (ADENOSINETRIPHOSPHATASE)  
 DE (ADENYLPHOSPHATASE) (ATP MONOPHOSPHATASE) (TRIPHOSPHATASE).  
 GN OSA1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIPPONBARE K-1;  
 RA WADA M., TAKANO M., KASAMO K.;  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + PHOSPHATE.  
 DR EMBL; D10207; BAA01058.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13241; Oryza; 494.1.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105079 MW; 1A9E387B CRC32;

Query Match 72.0%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 DB 61 KFLKFLGFM 69  
 ||||| |

RESULT 8

Q43002 ID Q43002 PRELIMINARY; PRT; 957 AA.  
 AC Q43002;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H<sup>+</sup>-ATPASE.  
 GN OSA2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RX MEDLINE; 95211383.  
 RA OOKURA T., WADA M., SAKAKIBARA Y., JEONG K.H., MARUTA I., KAWAMURA Y.,  
 RA KASAMO K.;  
 RT "Identification and characterization of a family of genes for the  
 RL plasma membrane H<sup>+</sup>-ATPase of Oryza sativa L.";  
 RL Plant Cell Physiol. 35:1251-1256(1994).  
 DR EMBL; D31843; BAA06629.1; -;  
 DR MENDEL; 13242; Oryza; 494.2.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 SQ SEQUENCE 957 AA; 105607 MW; 6673B764 CRC32;

Query Match 72.0%; Score 36; DB 10; Length 957;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLKFL 9  
 DB 61 KFLKFLGFM 69  
 ||||| |

RESULT 9  
 Q73742 ID Q73742 PRELIMINARY; PRT; 403 AA.  
 AC Q73742;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE XOE-2 (FRAGMENT).  
 GN XOE-2.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;  
 OC Xenopus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEAD;  
 RA BALDESSARI D., CONSULEZ G.G.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF040994; AAC15659.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 403 AA; 43800 MW; BBB9FC22 CRC32;

Query Match 72.0%; Score 36; DB 13; Length 403;  
 Best Local Similarity 77.8%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLKFLKFLK 10  
 DB 376 FLKFLKFKV 384  
 ||||| |

RESULT 10  
 Q26046 ID Q26046 PRELIMINARY; PRT; 1279 AA.  
 AC Q26046;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DN TYPE IIS RESTRICTION ENZYME R AND M PROTEIN (ECO57IR).  
 GN HP1517.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria: Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695;  
 RX MEDLINE: 97394467.  
 RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
 RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
 RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,  
 RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATHEY L., WALLIN E.,  
 RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the gastric pathogen Helicobacter  
 RT pylori."  
 RL Nature 388:539-547(1997).  
 DR EMBL: AF000849; AAD08556.1; -.  
 DR TIGR: HP1517; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1279 AA; 149714 MW; 074DBDA7 CRC32;

Query Match 70.08; Score 35; DB 2; Length 1279;  
 Best Local Similarity 80.08; Pred. No. 2.3e+02;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFLKFLKFLK 10  
 :|||:||||  
 Db 738 KFLKFLKFLK 747

RESULT 11  
 ID P78815 PRELIMINARY; PRT; 290 AA.  
 AC P78815;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE FISSION YEAST (FRAGMENT).  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PR745;  
 RX MEDLINE: 98162722.  
 RA YOSHIOKA S., KATO K., NAKAI K., OKAYAMA H., NOJIMA H.;  
 RT "Identification of open reading frames in Schizosaccharomyces pombe  
 RT conas."  
 RL DNA Res. 4:363-369(1997).  
 DR EMBL: D89164; BAA13826.1; -.  
 DR PFAM: PF00018; SH3; 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 290 AA; 33434 MW; 86761774 CRC32;

Query Match 70.0%; Score 35; DB 3; Length 290;  
 Best Local Similarity 66.7%; Pred. No. 66;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKFLKFLK 10  
 :|||:||||

Db 105 YLRFLOFLK 113

RESULT 12  
 Q9ZYM7  
 ID Q9ZYM7 PRELIMINARY; PRT; 552 AA.  
 AC Q9ZYM7;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE NADH DEHYDROGENASE 5.  
 OS Rhinoceros squalidus sanguineus.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodidae; Ixodidae; Rhinoceros.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 99083443.  
 RA BLACK W.C. IV, ROHRDANZ R.L.;  
 RT "Mitochondrial gene order is not conserved in arthropods: prostriate  
 RT and metastriate tick mitochondrial genomes.";  
 RL Mol. Biol. Evol. 15:1772-1785(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA BLACK W.C. IV, ROHRDANZ R.L.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF081829; AAD05525.1; -.  
 KW Mitochondrion.  
 SQ SEQUENCE 552 AA; 64542 MW; 82C26959 CRC32;

Query Match 70.0%; Score 35; DB 8; Length 552;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFLKFLKFLK 10  
 :|||:||||  
 Db 465 KFLKFLKFLK 474

RESULT 13  
 ID Q27551 PRELIMINARY; PRT; 260 AA.  
 AC Q27551;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE 2-HYDROXYHEPTA-2,4-DIENE-1,7-DIOATE ISOMERASE.  
 GN MTH1507.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE: 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,  
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL: AF000911; AAB85982.1; -.  
 KW Isomerase.  
 SQ SEQUENCE 260 AA; 28873 MW; 536811AB CRC32;

Query Match 68.0%; Score 34; DB 1; Length 260;  
 Best Local Similarity 75.0%; Pred. No. 87;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LKFLKFLK 10

Db 4 LKFLFKFLK 11  
|||||:|:|  
Best Local Similarity 87.5%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 FLKFLKFL 9  
Db 187 FLKFKKFL 194  
Search completed: September 7, 1999, 20:34:38  
Job time: 19734 sec

RESULT 14  
OS8773 PRELIMINARY; PRT; 132 AA.  
AC OS8773;  
DT 01-AUG-1998 (TRENBLrel. 07, Created)  
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
DE 132AA LONG HYPOTHETICAL PROTEIN.  
GN PH1024.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3.  
RX MEDLINE: 98344137.  
RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,  
RA SAKAI M., OGURA K., OTURA K., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
RA KIKUCHI H.;  
RT "Complete sequence and gene organization of the genome of a  
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL: AP000004; BAA30121.1; .  
SQ SEQUENCE 132 AA; 15657 MW; B29A2937 CRC32;

Query Match 68.0%; Score 34; DB 1; Length 132;  
Best Local Similarity 70.0%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFLKFLKFLK 10  
| ||| |||  
Db 121 KMLKPKFLR 130

RESULT 15  
OS1263 PRELIMINARY; PRT; 201 AA.  
AC OS1263;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE CONSERVED HYPOTHETICAL PROTEIN.  
GN BB0247.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE: 98065943.  
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,  
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
RA SMITH H.O., VENTER J.C.;  
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
burgdorferi.";  
RL Nature 390:580-586(1997).  
DR EMBL: AE001135; AAB91499.1; .  
DR TIGR: BB0247; .  
SQ SEQUENCE 201 AA; 23138 MW; 1F87F075 CRC32;

Query Match 68.0%; Score 34; DB 2; Length 201;

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:11 ; Search time 147.16 Seconds  
(without alignments)  
1.449 Million cell updates/sec

Title: US-09-124-280A-9

Perfect score: 48

Sequence: 1 RYRVYRVYV 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	1 R3534	Peptide for treati
2	48	100.0	9	1 R39296	Endotoxin lipid A
3	48	100.0	9	1 R71780	Peptide neutralisi
4	48	100.0	9	1 W21627	Antibiotic potenti
5	48	100.0	9	1 W21597	Antibiotic potenti
6	36	75.0	15	1 W12306	Immunoglobulin-bi
7	35	72.9	463	1 W82250	Mouse PTP05 isofor
8	35	72.9	405	1 W82251	Mouse PTP05 isofor
9	35	72.9	122	1 W82252	Rat PTP10. New nuc
10	35	72.9	426	1 W82249	Mouse PTP05. New n
11	34	70.8	394	1 W83113	O antigen polymera
12	32	66.7	6	1 R37386	Peptide for treati
13	31	64.6	1422	1 R32070	Hepatitis GB virus
14	30	62.5	553	1 R13944	Peptide for treati
15	30	62.5	6	1 R37387	Partial HVT ribonu
16	30	62.5	1013	1 R62535	P. vulgaris chondr
17	30	62.5	990	1 R84660	Chondroitinase II.
18	30	62.5	990	1 R77460	Chondroitinase-II.
19	30	62.5	568	1 W03325	Newcastle disease
20	30	62.5	568	1 W03138	Newcastle disease
21	30	62.5	1013	1 W09398	Proteus vulgaris c
22	30	62.5	568	1 W03551	Newcastle disease
23	30	62.5	990	1 W39746	P. vulgaris chondr
24	30	62.5	990	1 W37786	Amino acid sequenc
25	30	62.5	144	1 W55084	Streptococcus pneu
26	30	62.5	568	1 W62989	HN protein of Newc
27	30	62.5	571	1 W44940	Newcastle disease
28	30	62.5	568	1 W49683	Newcastle disease
29	30	62.5	30	1 W74846	Human secreted pro
30	30	62.5	1118	1 W82395	Human UB protein
31	30	62.5	990	1 W90075	P. vulgaris chondr
32	30	62.5	194	1 Y11220	S. pneumoniae prot
33	29	60.4	448	1 R14478	51.4 kD toxin, Cau
34	29	60.4	554	1 R26271	Alpha-acetolactate
35	29	60.4	241	1 R29515	Caffeoyl-CoA-3-O-m
36	29	60.4	448	1 R41016	Insecticidal prote
37	29	60.4	448	1 R49873	Bacillus sphaericu
38	29	60.4	377	1 R56529	Protein kinase (CK
39	29	60.4	347	1 R56530	Human HRR25-like c
40	29	60.4	377	1 R75625	Human HRR25-like c
41	29	60.4	347	1 R75626	Human HRR25-like c
42	29	60.4	98	1 W55298	H. pylori ORF 06cp
43	29	60.4	134	1 W55658	H. pylori ORF 11ae

ALIGNMENTS

RESULT 1

R3534  
ID R3534 standard; peptide; 9 AA.  
AC R3534;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 12; Page 32; 39pp; English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

RESULT 2

R39296  
ID R39296 standard; peptide; 9 AA.  
AC R39296;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
KW quantification.  
OS Synthetic.  
PN W09314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.

PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 12; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9  
 | | | | | | | |  
 Db 1 RYRVYRVYV 9

## RESULT 3

R71780  
 ID R71780 standard; peptide: 9 AA.  
 AC R71780;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 11; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9  
 | | | | | | | |  
 Db 1 RYRVYRVYV 9

## RESULT 4

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

W21627  
 ID W21627 standard; peptide; 9 AA.  
 AC W21627;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #39.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 43; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9  
 | | | | | | | |  
 Db 1 RYRVYRVYV 9

## RESULT 5

W21597  
 ID W21597 standard; peptide; 9 AA.  
 AC W21597;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #9.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 13; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRYRYRYV 9  
 Db 1 RYRYRYRYV 9

RESULT 6  
 W12306  
 ID W12306 standard; peptide; 15 AA.  
 AC W12306;  
 DT 29-AUG-1997 (first entry)  
 DE Immunoglobulin-binding peptide used as affinity and capture ligand.  
 KW Immunoglobulin binding; affinity chromatography; immunoassay;  
 KW capture ligand; constant region; non-covalent bond.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 1. .12  
 FT /note= "optionally D form residue"  
 PN EP-752425-A2.  
 PD 08-JAN-1997.  
 PF 19-JUN-1996; 201706.  
 PR 21-JUN-1995; IT-M1328.  
 PR 29-APR-1996; IT-M10831.  
 PA (TECN-) TECNOGEN SCPA.  
 PI Fassina G, Ruvo M, Verdoliva A;  
 DR WPI: 97-067416/07.  
 PT New immunoglobulin-binding peptide(s) - for use in affinity  
 PT chromatography and immunoassay  
 PS Example 2; Page 7; 12pp; English.  
 CC W12305-W12306 are peptides that non-covalently bind to the constant  
 CC region of an immunoglobulin. They are used as affinity ligands for  
 CC separation and purification of immunoglobulins (Ig), and as capture  
 CC ligands for binding Ig to microtitre plates for ELISA determination.  
 SQ Sequence 15 AA;

Query Match 75.0%; Score 36; DB 1; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 0.23;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRYRYRY 8  
 Db 3 RYRYRYRY 10

RESULT 7  
 W89250  
 ID W89250 standard; Protein; 463 AA.  
 AC W89250;  
 DT 10-MAR-1999 (first entry)  
 DE Mouse PTP05 isoform #1.  
 KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;  
 KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;  
 KW neurodegenerative disease; neuronal survival; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease.  
 OS Mus sp.  
 PN WO9849317-A2.  
 PD 05-NOV-1998.  
 PF 27-APR-1998; U08439.  
 PR 23-OCT-1997; US-063595.  
 PR 28-APR-1997; US-044428.  
 PR 20-MAY-1997; US-047222.  
 PR 11-JUN-1997; US-049477.  
 PR 11-JUN-1997; US-049756.  
 PR 18-JUN-1997; US-049914.  
 PA (SUGEN-) SUGEN INC.  
 PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,  
 PI Markby D, Onrust S, Peles E, Plozman GD;  
 DR WPI: 99-009434/01.  
 DR N-PSDB; V81745.

PT New nucleic acid encoding specific protein tyrosine phosphatases -  
 PT useful for identifying specific modulators for treatment and  
 PT prevention of cancer and neurodegenerative disease

PS Claim 2; Page 157-158; 193pp; English.  
 CC The present invention describes isolated, enriched or purified nucleic  
 CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The  
 CC present sequence represents mouse PTP05. The above proteins, other than  
 CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify  
 CC substances that modulate their activity (i.e. agonists and antagonists,  
 CC including NBP) in vivo or in vitro. These substances are used to treat  
 CC or prevent diseases associated with abnormal signal transduction  
 CC pathways that involve the proteins, particularly cancer (e.g. leukaemia  
 CC and lymphoma), while modulators of ALK-7 (which is a type I receptor  
 CC serine/threonine kinase) are used to promote neuronal survival,  
 CC particularly for treating Alzheimer's, Parkinson's or Huntington's  
 CC diseases. Nucleic acid fragments of the polynucleotides encoding the  
 CC proteins can be used as probes to identify and clone related sequences;  
 CC to detect protein-encoded RNA; to generate transgenic animals and in  
 CC gene therapy (optionally after mutation). Ab are used to determine the  
 CC proteins.  
 SQ Sequence 463 AA;

Query Match 72.9%; Score 35; DB 1; Length 463;  
 Best Local Similarity 62.5%; Pred. No. 16;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRYRYRYV 9  
 Db 376 FIKYRYRYV 383

RESULT 8  
 W89251  
 ID W89251 standard; Protein; 405 AA.  
 AC W89251;  
 DT 10-MAR-1999 (first entry)  
 DE Mouse PTP05 isoform #2.  
 KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;  
 KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;  
 KW neurodegenerative disease; neuronal survival; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease.  
 OS Mus sp.  
 PN WO9849317-A2.  
 PD 05-NOV-1998.  
 PF 27-APR-1998; U08439.  
 PR 23-OCT-1997; US-063595.  
 PR 28-APR-1997; US-044428.  
 PR 20-MAY-1997; US-047222.  
 PR 11-JUN-1997; US-049477.  
 PR 11-JUN-1997; US-049756.  
 PR 18-JUN-1997; US-049914.  
 PA (SUGEN-) SUGEN INC.  
 PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,  
 PI Markby D, Onrust S, Peles E, Plozman GD;  
 DR WPI: 99-009434/01.  
 DR N-PSDB; V81746.

PT New nucleic acid encoding specific protein tyrosine phosphatases -  
 PT useful for identifying specific modulators for treatment and  
 PT prevention of cancer and neurodegenerative disease

PS Claim 2; Page 158-160; 193pp; English.  
 CC The present invention describes isolated, enriched or purified nucleic  
 CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The  
 CC present sequence represents mouse PTP05. The above proteins, other than  
 CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify  
 CC substances that modulate their activity (i.e. agonists and antagonists,  
 CC including NBP) in vivo or in vitro. These substances are used to treat  
 CC or prevent diseases associated with abnormal signal transduction  
 CC pathways that involve the proteins, particularly cancer (e.g. leukaemia  
 CC and lymphoma), while modulators of ALK-7 (which is a type I receptor  
 CC serine/threonine kinase) are used to promote neuronal survival,  
 CC particularly for treating Alzheimer's, Parkinson's or Huntington's  
 CC diseases. Nucleic acid fragments of the polynucleotides encoding the  
 CC proteins can be used as probes to identify and clone related sequences;  
 CC to detect protein-encoded RNA; to generate transgenic animals and in  
 CC gene therapy (optionally after mutation). Ab are used to determine the

```

CC proteins.
SQ Sequence 405 AA;

Query Match 72.9%; Score 35; DB 1; Length 405;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRVRYV 9
   ::::|
Db 339 FIKVRYV 346

RESULT 9
W89252
ID W89252 standard; Protein; 122 AA.
AC W89252;
DT 10-MAR-1999 (first entry)
DE Rat PTP10.
KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
KW neurodegenerative disease; neuronal survival; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease.
OS Rattus sp.
PN WO9849317-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08439.
PR 23-OCT-1997; US-063595.
PR 28-APR-1997; US-044428.
PR 20-MAY-1997; US-047222.
PR 11-JUN-1997; US-049477.
PR 11-JUN-1997; US-049756.
PR 18-JUN-1997; US-049914.
PA (SUGEN-) SUGEN INC.
PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,
PI Markby D, Onrust S, Peles E, Plowman GD;
DR WPI: 99-009434/01.
DR N-PSDB: V81747.
PT New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
PS Claim 2; Page 160; 193pp; English.
CC The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence represents rat PTP10. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC including NBP) in vivo or in vitro. These substances are used to treat
CC or prevent diseases associated with abnormal signal transduction
CC pathways that involve the proteins, particularly cancer (e.g. leukaemia
CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
SQ Sequence 122 AA;

Query Match 72.9%; Score 35; DB 1; Length 122;
Best Local Similarity 62.5%; Pred. No. 3.6;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRVRYV 9
   ::::|
Db 96 FIKVRYV 103

RESULT 10
W89249
ID W89249 standard; Protein; 426 AA.

```

```

AC W89249;
DT 10-MAR-1999 (first entry)
DE Mouse PTP05.
KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
KW neurodegenerative disease; neuronal survival; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease.
OS Mus sp.
PN WO9849317-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08439.
PR 23-OCT-1997; US-063595.
PR 28-APR-1997; US-044428.
PR 20-MAY-1997; US-047222.
PR 11-JUN-1997; US-049477.
PR 11-JUN-1997; US-049756.
PR 18-JUN-1997; US-049914.
PA (SUGEN-) SUGEN INC.
PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,
PI Markby D, Onrust S, Peles E, Plowman GD;
DR WPI: 99-009434/01.
DR N-PSDB: V81744.
PT New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
PS Claim 2; Page 155-157; 193pp; English.
CC The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence represents mouse PTP05. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC including NBP) in vivo or in vitro. These substances are used to treat
CC or prevent diseases associated with abnormal signal transduction
CC pathways that involve the proteins, particularly cancer (e.g. leukaemia
CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
SQ Sequence 426 AA;

Query Match 72.9%; Score 35; DB 1; Length 426;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRVRYV 9
   ::::|
Db 339 FIKVRYV 346

RESULT 11
W88313
ID W88313 standard; Protein; 394 AA.
AC W88313;
DT 26-APR-1999 (first entry)
DE O antigen polymerase involved in O157 antigen biosynthesis.
KW O antigen; O157 antigen; polymerase; wzy gene; diarrhoea;
KW haemorrhagic colitis; diagnosis.
OS Escherichia coli.
PN WO9850531-A1.
PD 12-NOV-1998.
PF 01-MAY-1998; AU0315.
PR 22-JUL-1997; AU-008162.
PR 01-MAY-1997; AU-006545.
PA (UNSW ) UNIV SYDNEY.
PI Reeves PR, Wang L.
DR WPI: 99-059669/05.
DR N-PSDB: X06749.
PT Nucleic acid molecules specific for bacterial polysaccharide

```

PT antigens - useful for detecting specific strains in, e.g. food,  
 PT faeces or patient samples  
 PS Disclosure; Fig 8; 165pp; English.  
 CC This is the amino acid sequence of an O antigen polymerase that is  
 CC encoded by the wzy open reading frame of a gene cluster (see  
 CC X06749) involved in the biosynthesis of the Escherichia coli O157  
 CC O antigen. The use of nucleic acid molecules derived from  
 CC particular assembly and transport genes, particularly wbd  
 CC (transferase), wzx (flippase) and wzy (polymerase) genes, within O  
 CC antigen gene clusters improves the specificity of methods for the  
 CC detection and identification of O antigens, e.g. in testing food-  
 CC or faecal-derived samples, or samples from patients. The O antigen  
 CC is a major virulence factor of enteropathogenic E. coli strains  
 CC that cause diarrhoea and haemorrhagic colitis.  
 SQ Sequence 394 AA;

Query Match 70.8%; Score 34; DB 1; Length 394;  
 Best Local Similarity 57.1%; Pred. No. 20;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRYRYR 8  
 Db 117 YRYRYR 123

## RESULT 12

ID R37386 standard; peptide; 6 AA.  
 AC R37386;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B; Group II.  
 OS Synthetic.

FH Key Location/Qualifiers

FT region 1..3

FT region 4..6

FT region /note= "repeat region"

FT region /note= "repeat region"

PN ZA9200943-A.

PD 25-NOV-1992; 000943.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORR M.

PI Porro M.

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated

PS Example; Page 5; 39pp; English.

CC The (Group II) peptide is an example of a generic peptide of formula  
 CC R-(Lys/Arg/His-Phe/Tyr/Trp-Leu/Ile/Val)n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock.

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

SQ Sequence 6 AA;

Query Match 66.7%; Score 32; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRYRV 6

Db 1 RYRYRV 6

## RESULT 13

ID R82070 standard; Protein; 1422 AA.

AC R82070;

DT 02-JUL-1996 (first entry)

DE Hepatitis GB virus (HGBV) clone protein prod.

KW Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine;

KW reagents; non-A; non-B; non-C; non-D; non-E; clone;

KW tamarin; infected plasma; lambda phage; cDNA library.

OS Hepatitis GB virus.

FH Key Location/Qualifiers

FT misc\_difference 1..1422

FT /note= "others correspond to degenerate or STOP

FT codons in T00052"

PN WO9521922-A2.

PD 17-AUG-1995.

PF 14-FEB-1995; U02118.

PR 14-FEB-1994; US-196030.

PR 13-MAY-1994; US-242654.

PR 29-JUL-1994; US-283314.

PR 23-NOV-1994; US-344190.

PR 23-NOV-1994; US-344185.

PR 27-JAN-1995; US-344557.

PA (ABSO) ABBOTT LAB.

PI Buijck SL, Dawson GJ, Desai SM, Erker JC, Leary TP;

PI Muerhoff AS, Mushahwar IK, Pilot-Matias TJ, Schlauder GG;

PI Simons JN;

DR WPI; 95-293123/38.

DR N-PSDB; T00052.

PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful

PT for diagnosis and therapy of hepatitis GB virus

PS Example 5; Pages 254-259; 661pp; English.

CC Double stranded hepatitis GB virus (HGBV) DNA obt'd. from HGBV

CC infected tamarin plasma, using standard procedures, was used to

CC prepare a lambda phage HGBV cDNA library. The cDNA clone T00052,

CC which encodes the proteins R82066-71 (the 6 possible reading

CC frames), was rescued from the lambda phage, searched against a

CC sequence database and found to be an unique HGBV sequence.

CC Reagents which comprise the HGBV DNA, or its protein prods. can

CC be used for the diagnosis, therapy or in a vaccine to prevent

CC HGBV infection.

SQ Sequence 1422 AA;

Query Match 64.6%; Score 31; DB 1; Length 1422;

Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRYRYRV 9

Db 1213 RYMPYVLYV 1221

## RESULT 14

ID R13944 standard; Protein; 553 AA.

AC R13944;

DT 27-NOV-1991 (first entry)

DE Partial HVT ribonucleotide reductase large subunit RRL.

KW Newcastle disease; recombinant poultry virus vaccine.

OS Turkey herpes virus strain FC 126.

PN EP-447303-A.

PD 18-SEP-1991.

PF 07-MAR-1991; 400634.

PR 12-MAR-1990; FR-003105.

PA (RHON-) RHONE MERIEUX SA.

PI Rev-Senelonge A, Kohen G;  
DR WPI: 91-275886/38.

DR N-PSDB; Q13430.

PT New herpes recombinants - useful as vaccines against human and  
PT animal viral conditions, e.g. fowl pest, coccidiosis and  
PT Pasteurellosis

PS Disclosure: Page 13-14; 22pp; French.

CC This sequence is the C-terminal region of HVT RR1. It was deduced  
CC from a genomic DNA clone which was isolated because it contained  
CC the entire RR2 gene. Knowledge of the RR2 small sub-unit sequence is  
CC exploited in the construction of recombinant vaccines. A heterologous  
CC sequence coding for an appropriate immunogen to protect against e.g.  
CC Newcastle disease, avian anaemia, colibacillosis, avian infectious  
CC bronchitis, etc. is inserted into the RR2 gene. The HVT is suitable  
CC for use in live vaccines as it is apathogenic and non-oncogenic.  
CC See also Q13668.

SQ Sequence 553 AA;

Query Match 62.5%; Score 30; DB 1; Length 553;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRYV 9

Db 108 FKRYLYRV 115

Y : ||:||||

RESULT 15

R37387

ID R37387 standard; peptide; 6 AA.

AC R37387;

DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;

KW cytokine release control; treatment; pertussis; bacterial meningitis;

KW HIV related infections; polymyxin B; Group II.

OS Synthetic.

FH Key Location/Qualifiers

FT region 1. .3

FT /note= "repeat region"

FT region 4. .6

FT /note= "repeat region"

FN Z9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M;

DR WPI: 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises

PT specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated

PS Example; Page 5; 39pp; English.

CC The (Group II) peptide is an example of a generic peptide of formula

CC R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock,

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

SQ Sequence 6 AA;

Query Match

Best Local Similarity 62.5%; Score 30; DB 1; Length 6;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRYRV 6

Db 1 RYIRYI 6

Y : ||:||||

Search completed: September 7, 1999, 20:37:12  
Job time: 18461 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:34 ; Search time 80.79 Seconds  
(without alignments)  
1.099 Million cell updates/sec

Title: US-09-124-280A-9  
Perfect score: 48  
Sequence: 1 RYRVRYRV 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	1	US-07-819-893-10
2	48	100.0	9	1	US-08-280-397-10
3	48	100.0	9	1	US-08-097-830E-9
4	48	100.0	9	2	US-08-456-112B-9
5	48	100.0	9	2	US-08-456-112B-39
6	33	68.8	9	1	US-08-097-830E-31
7	30	62.5	990	1	US-08-232-540-2
8	30	62.5	1013	1	US-08-233-008A-8
9	30	62.5	990	1	US-08-428-949A-2
10	30	62.5	571	1	US-08-368-803-17
11	30	62.5	990	1	US-08-428-948A-2
12	30	62.5	990	1	US-08-428-946-2
13	30	62.5	200	2	US-08-606-143-42
14	30	62.5	990	3	PCT-US95-04656-2
15	29	60.4	554	1	US-07-839-433-2
16	29	60.4	377	1	US-08-454-097-31
17	29	60.4	347	1	US-08-454-097-33
18	29	60.4	448	2	US-08-231-342-23
19	28	58.3	287	1	US-08-457-245-3
20	27	56.2	577	1	US-07-820-154A-30
21	27	56.2	1167	1	US-08-485-568A-6
22	27	56.2	1168	1	US-08-620-717A-9
23	27	56.2	599	2	US-08-752-238-3
24	27	56.2	1167	2	US-08-590-554A-6
25	27	56.2	536	2	US-08-551-211-3
26	27	56.2	577	2	US-08-663-566A-11
27	27	56.2	577	2	US-08-097-554A-30
28	27	56.2	2254	2	US-08-286-819A-28
29	27	56.2	788	2	US-08-918-914-4
30	27	56.2	457	2	US-08-882-704A-6
31	27	56.2	577	3	PCT-US93-0034-30
32	27	56.2	140	3	PCT-US93-11612-4
33	27	56.2	121	3	PCT-US93-11612-7
34	27	56.2	121	3	PCT-US93-11612-8
35	27	56.2	140	3	PCT-US93-11612-12
36	27	56.2	591	3	PCT-US94-01826A-13
37	27	56.2	577	3	PCT-US94-02252A-13
38	27	56.2	577	3	PCT-US95-10245-11
39	27	56.2	78	4	5432081-9

ALIGNMENTS

RESULT 1  
US-07-819-893-10  
; Sequence 10, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: AMINO ACIDS  
; TOPOLOGY: linear  
US-07-819-893-10

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04; Indels 0;  
Matches 9; Conservative 0; Mismatches 0; Gaps 0;

QY 1 RYRVRYRV 9  
111111111  
Db 1 RYRVRYRV 9

RESULT 2  
US-08-280-397-10  
; Sequence 10, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acids  
;; TOPOLOGY: linear  
US-08-280-397-10

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
DB 1 RYRVRYRV 9

RESULT 3  
US-08-097-830E-9  
;; Sequence 9, Application US/08097830E  
;; Patent No. 5652211  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Peptides For Neutralizing The  
;; TITLE OF INVENTION: Toxicity of Lipid A  
;; NUMBER OF SEQUENCES: 35  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/097,830E  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:

;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
US-08-097-830E-9

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
DB 1 RYRVRYRV 9

RESULT 4  
US-08-456-112B-9  
;; Sequence 9, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
US-08-456-112B-9

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
DB 1 RYRVRYRV 9



Db 1 RYRVYRVYV 9

## RESULT 5

US-08-456-112B-39  
Sequence 39, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular

US-08-456-112B-39

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

## RESULT 6

US-08-097-830E-31  
Sequence 31, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular

US-08-097-830E-31

Query Match 68.8%; Score 33; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRVYRVYV 9

Db 2 YVEIVEIV 9

## RESULT 7

US-08-232-540-2  
Sequence 2, Application US/08232540  
Patent No. 5498536  
GENERAL INFORMATION:  
APPLICANT: Khandke, Kiran M.  
TITLE OF INVENTION: A No. 5498536el Protein Designated  
Chondroitinase II and its Use With A Protein Designated  
Chondroitinase I to Achieve Complete Vitreal Disinsertion  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,540  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Gordon, Alan M.  
REGISTRATION NUMBER: 30,637  
REFERENCE/DOCKET NUMBER: 32,390  
TELEPHONE: 201-831-3244  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-232-540-2

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9

Db 1 RYRVYRVYV 9

Query Match 62.5%; Score 30; DB 1; Length 990;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYVRY 8  
 |||||:  
 Db 615 RYGRYLQY 622

## RESULT 8

US-08-233-008A-8  
 ; Sequence 8, Application US/082333008A  
 ; Patent No. 5578480  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Khandke, Kiran M.  
 ; TITLE OF INVENTION: Methods For The Isolation And  
 ; Purification Of The Recombinantly Expressed chondroitinase  
 ; TITLE OF INVENTION: I and II Enzymes From P. vulgaris  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: American Cyanamid Company  
 ; STREET: One Cyanamid Plaza  
 ; CITY: Wayne  
 ; STATE: New Jersey  
 ; COUNTRY: U.S.A.  
 ; ZIP: 07470-8426  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/233,008A  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gordon, Alan M.  
 ; REGISTRATION NUMBER: 30,637  
 ; REFERENCE/DOCKET NUMBER: 31,885-01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-831-3244  
 ; TELEFAX: 201-831-3305  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1013 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-233-008A-8

Query Match 62.5%; Score 30; DB 1; Length 1013;  
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYVRY 8  
 |||||:  
 Db 638 RYGRYLQY 645

## RESULT 9

US-08-428-949A-2  
 ; Sequence 2, Application US/08428949A  
 ; Patent No. 5716617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kiran M. Khandke, John Gotto, Ursula Eul  
 ; TITLE OF INVENTION: Compositions of Chondroitinase I and  
 ; TITLE OF INVENTION: Chondroitinase II  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Darby & Darby P.C.

STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022-7513  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/428,949A  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Robinson, Joseph R.  
 ; REGISTRATION NUMBER: 33,448  
 ; REFERENCE/DOCKET NUMBER: 0646/1B017-US2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-527-7783  
 ; TELEFAX: 212-753-6237  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 990 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-428-949A-2

## Query Match

62.5%; Score 30; DB 1; Length 990;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYVRY 8  
 |||||:  
 Db 615 RYGRYLQY 622

## RESULT 10

US-08-368-803-17  
 ; Sequence 17, Application US/08368803  
 ; Patent No. 5733554  
 ; GENERAL INFORMATION:  
 ; APPLICANT: AUDONNET, Jean-Christophe F  
 ; APPLICANT: BUBLOT, Michel J  
 ; APPLICANT: DARTEIL, Raphael J  
 ; APPLICANT: DUINAT, Carole V  
 ; APPLICANT: LAPLACE, Eliane L  
 ; APPLICANT: RIVIERE, Michel A  
 ; TITLE OF INVENTION: Avian Herpesvirus-based live recombinant avian  
 ; TITLE OF INVENTION: vaccine, in particular against Gumboro disease  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LARSON AND TAYLOR  
 ; STREET: 727 SOUTH 23RD STREET  
 ; CITY: ARLINGTON  
 ; STATE: VIRGINIA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/368,803  
 ; FILING DATE: 05-JAN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: SARRO, Thomas  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 920-7200

TELEFAX: (703) 892-8428  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 571 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-368-803-17

Query Match 62.5%; Score 30; DB 1; Length 571;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVRYVRY 8  
||| |||  
Db 333 RYVRYVRY 340

RESULT 11  
US-08-428-948A-2  
Sequence 2, Application US/08428948A  
Patent No. 5741692  
GENERAL INFORMATION:  
APPLICANT: Khandke, Kiran M.  
TITLE OF INVENTION: Chondroitinase II  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,948A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Robinson, Joseph R.  
REGISTRATION NUMBER: 33,448  
REFERENCE/DOCKET NUMBER: 0646/1B017-US4  
TELEPHONE: 212-527-7783  
TELEFAX: 212-753-6327  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-948A-2

Query Match 62.5%; Score 30; DB 1; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYVRY 8  
||| |||  
Db 615 RYGRYLOY 622

RESULT 12  
US-08-428-946-2  
Sequence 2, Application US/08428946  
Patent No. 5855883  
GENERAL INFORMATION:

APPLICANT: Kiran M. Khandke, John Gotto, Ursula Eul  
TITLE OF INVENTION: Method of Disinsertion of Vitreous Body from  
NEURAL RETINA OF THE EYE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,946  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Robinson, Joseph R.  
REGISTRATION NUMBER: 33,448  
REFERENCE/DOCKET NUMBER: 0646/1B017-US3  
TELEPHONE: 212-527-7783  
TELEFAX: 212-753-6327  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-946-2

Query Match 62.5%; Score 30; DB 2; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYVRY 8  
||| |||  
Db 615 RYGRYLOY 622

RESULT 13  
US-08-606-143-42  
Sequence 42, Application US/08606143  
Patent No. 5856155  
GENERAL INFORMATION:  
APPLICANT: Li, Min  
TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR  
MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH  
TITLE OF INVENTION: COMPOUNDS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Volt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,143  
FILING DATE: 23-FEB-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kilyk Jr., John

```

; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-606-143-42

```

```

Query Match          62.5%; Score 30; DB 2; Length 200;
Best Local Similarity 62.5%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 RYRVRY 8
   |||||
Db 181 RYRVYAF 188

```

```

RESULT 14
PCT-US95-04656-2
; Sequence 2, Application PC/TUS9504656
; GENERAL INFORMATION:
; APPLICANT: Khandke, Kiran M.
; TITLE OF INVENTION: A Novel Protein Designated
; TITLE OF INVENTION: Chondroitinase II and its Use With a Protein Designated
; TITLE OF INVENTION: Chondroitinase I to Achieve Complete Vitreal Disinsertion
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04656
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gordon, Alan M.
; REGISTRATION NUMBER: 30,637
; REFERENCE/DOCKET NUMBER: 32,390-00/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-831-3244
; TELEFAX: 201-831-3305
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 990 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04656-2

```

```

Query Match          62.5%; Score 30; DB 3; Length 990;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 RYRVRY 8
   |||||
Db 615 RYGRLOY 522

```

```

RESULT 15
US-07-839-433-2
; Sequence 2, Application US/07839433
; Patent No. 5420021
; GENERAL INFORMATION:
; APPLICANT: MARUGG, JOHN D
; APPLICANT: TOONEN, MARIA Y
; APPLICANT: VERHUE, WALTER M
; APPLICANT: VERRIES, CORNELIS T
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF
; TITLE OF INVENTION: ALPHA-ACETOLACTIC ACID
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARRY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/839,433
; FILING DATE: 19920224
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 94713
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-839-433-2

```

```

Query Match          60.4%; Score 29; DB 1; Length 554;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 YRVRY 8
   |||||
Db 493 YVDYVY 499

```

```

Search completed: September 7, 1999, 22:38:34
Job time: 7918 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:23 ; Search time 116.8 Seconds  
(without alignments)  
3.087 Million cell updates/sec

Title: us-09-124-280A-9  
Perfect score: 48  
Sequence: 1 RYRVRYV 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	35	72.9	110	1 S73233	ribosomal protein
2	34	70.8	261	2 H69217	hypothetical prote
3	34	70.8	675	2 D48492	kpsc protein - Esc
4	34	70.8	600	2 S48509	transcription regu
5	33	68.8	1294	2 S77690	probable membrane
6	33	68.8	285	2 C65012	hypothetical prote
7	33	68.8	113	2 F64001	hypothetical prote
8	33	68.8	630	2 C39930	hypothetical prote
9	32	66.7	129	2 S53699	ribosomal protein
10	32	66.7	638	2 G02068	white homolog - hu
11	32	66.7	322	2 A70189	hypothetical prote
12	32	66.7	400	2 S32879	lipA protein - Nei
13	32	66.7	481	2 H70579	probable membrane
14	32	66.7	208	2 C64347	hypothetical prote
15	32	66.7	502	2 S26004	18S rRNA intron 1
16	32	66.7	785	2 S63652	hypothetical prote
17	32	66.7	507	2 S52648	myo-inositol-1-pho
18	31	64.6	813	1 A49123	fibroblast growth
19	31	64.6	391	2 S69192	serine O-acetyltra
20	31	64.6	480	2 A56182	fibroblast growth
21	31	64.6	480	2 B56182	fibroblast growth
22	31	64.6	170	2 D65001	hypothetical prote
23	31	64.6	165	2 C2465	beta protein homol
24	31	64.6	346	2 E70791	probable lyase - M
25	31	64.6	1259	2 S25954	gene atpA intron 2
26	31	64.6	239	2 S35315	cdc16 protein - fi
27	31	64.6	514	2 C70446	hypothetical prote
28	31	64.6	223	2 B71057	hypothetical prote
29	30	62.5	1603	1 BVAS41	arom protein - Eme
30	30	62.5	570	1 HNNZAV	hemagglutinin-neur
31	30	62.5	571	1 H46328	hemagglutinin-neur
32	30	62.5	571	1 I46328	hemagglutinin-neur
33	30	62.5	571	1 A36829	hemagglutinin-neur
34	30	62.5	571	1 B36829	hemagglutinin-neur
35	30	62.5	571	2 S07126	hemagglutinin-neur
36	30	62.5	571	2 S40164	hemagglutinin-neur
37	30	62.5	581	2 S45114	hemagglutinin-neur
38	30	62.5	436	2 S66171	ribulose-bisphosph
39	30	62.5	228	2 C69859	two-component resp

ALIGNMENTS

RESULT 1

S73233

ribosomal protein L23 - red alga (Porphyra purpurea) chloroplast

C:Species: chloroplast Porphyra purpurea

C:Date: 23-Oct-1998 #sequence\_revision 23-Oct-1998 #text\_change 23-Oct-1998

C:Accession: S73233

R:Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A:Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A:Reference number: S73108

A:Accession: S73233

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-110 <REI>

A:Cross-references: EMBL:U38804; NID:g1276652; PID:g1276778

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: rpl23

A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein L23

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 72.9%; Score 35; DB 1; Length 110;

Best Local Similarity 55.0%; Pred. NO. 4.6;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVRYV 9

||:|:|:|:

DB 2 RYLKRYI 10

RESULT 2

H69217

hypothetical protein MTH882 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 03-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Sep-1998

C:Accession: H69217

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T

; Qiu, D.; Spadafora, R.; Viceire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,

kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: H69217

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <MTH>

A:Cross-references: GB:AE000864; GB:AE000666; NID:g2621970; PID:g2621976

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH882

A:Start codon: TTG

C:Superfamily: hypothetical protein HI0902

Query Match 70.8%; Score 34; DB 2; Length 261;

Best Local Similarity 71.4%; Pred. No. 17;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYRVRYV 9

||:|:|:|:

Db 241 VYIRYI 247

## RESULT 3

D48492  
kpsc protein - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 09-Sep-1997  
C:Accession: D48492; I67694; S36652  
R:Pazzani, C.; Rosenow, C.; Boulnois, G.J.; Bronner, D.; Jann, K.; Roberts, I.S.  
J. Bacteriol. 175, 5978-5983, 1993  
A:Title: Molecular analysis of region 1 of the Escherichia coli K5 antigen gene cluster.  
A:Reference number: A48492; MUID:93388530  
A:Accession: D48492  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-675 <PAZ>  
A:Cross-references: EMBL:X74567; NID:9397404; PID:9397408  
R:Rosenow, C.; Roberts, I.S.; Jann, K.  
FEMS Microbiol. Lett. 125, 159-164, 1995  
A:Title: Isolation from recombinant Escherichia coli and characterization of CMP-Kdo synthetase from Escherichia coli K5.  
A:Reference number: I53591; MUID:95180691  
A:Accession: I67694  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-8 <RES>  
A:Cross-references: GB:S76943; NID:9913365

Query Match 70.8%; Score 34; DB 2; Length 675;

Best Local Similarity 62.5%; Pred. No. 43;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVYRVYV 9

I:|:| |:

Db 293 YLYRCYI 300

## RESULT 4

S48509  
transcription regulator SP08 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein G1337; protein YGL192W  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 06-Feb-1998  
C:Accession: S48509; S62053; S64209  
R:Sakurai, M.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: S48509  
A:Accession: S48509  
A:Molecule type: DNA  
A:Residues: 1-600 <SAK>  
A:Cross-references: EMBL:D23721; NID:9433221; PID:dl005476; PID:g471168  
R:Coglievina, M.; Delneri, D.; Zaccaria, P.; Klima, R.; Bertani, I.; Bruschi, C.V.  
submitted to the EMBL Data Library, September 1995  
A:Description: A 6.7 kb fragment from chromosome VII of Saccharomyces cerevisiae contain  
A:Reference number: S62051  
A:Accession: S62053  
A:Molecule type: DNA  
A:Residues: 1-600 <COG>  
A:Cross-references: EMBL:X91837; NID:g1177627; PID:e203623; PID:g1177636  
A:Experimental source: strain FY1679  
R:Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64183  
A:Accession: S64209  
A:Molecule type: DNA  
A:Residues: 1-600 <BRU>  
A:Cross-references: EMBL:Z72714; NID:g1322814; PID:e243795; PID:g1322815; MIPS:YGL192W  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:IME4; SP08  
A:Cross-references: SGD:S0003160; MIPS:YGL192W  
A:Map position: 7L

C:Keywords: nucleus; transcription regulation

Query Match 70.8%; Score 34; DB 2; Length 600;  
Best Local Similarity 44.4%; Pred. No. 39;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYRVYV 9

I:|:| |:

Db 272 RYHYLYQYI 280

## RESULT 5

S77690  
Probable membrane protein YOL075c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein Oll125; hypothetical protein Oll130; hypothetical  
C:Species: Saccharomyces cerevisiae  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Jul-1998  
C:Accession: S77690; S66767; S66768  
R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66756  
A:Accession: S77690  
A:Molecule type: DNA  
A:Residues: 1-1294 <ALE>  
A:Cross-references: EMBL:Z74816; MIPS:YOL075c  
A:Note: this is a revision to the sequence from reference S66756  
R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66756  
A:Accession: S66767  
A:Molecule type: DNA  
A:Residues: 1-179, 'TTPTGVLVVKRED' <ALW>  
A:Cross-references: EMBL:Z74816  
A:Experimental source: strain S288C  
A:Note: this sequence has been revised in reference S77690  
A:Accession: S66768  
A:Molecule type: DNA  
A:Residues: 200-1294 <ALF>  
A:Cross-references: EMBL:Z74817  
A:Experimental source: strain S288C  
A:Note: this sequence has been revised in reference S77690  
C:Genetics:  
A:Map position: 15L  
A:Note: YOL075c

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog  
C:Keywords: P-loop; transmembrane protein  
F:45-263/Domain: ATP-binding cassette homology <ABC1>  
F:62-69/Region: nucleotide-binding motif A (P-loop)  
F:376-392/Domain: transmembrane #status predicted <TM1>  
F:469-485/Domain: transmembrane #status predicted <TM2>  
F:496-512/Domain: transmembrane #status predicted <TM3>  
F:606-622/Domain: transmembrane #status predicted <TM4>  
F:710-916/Domain: ATP-binding cassette homology <ABC2>  
F:727-734/Region: nucleotide-binding motif A (P-loop)  
F:1042-1058/Domain: transmembrane #status predicted <TM5>  
F:1125-1141/Domain: transmembrane #status predicted <TM6>  
F:1177-1193/Domain: transmembrane #status predicted <TM7>  
F:1269-1285/Domain: transmembrane #status predicted <TM8>

Query Match 68.8%; Score 33; DB 2; Length 1294;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVYRVYV 9

I:|:| |:

Db 550 YVYRWIKYI 557

## RESULT 6

## C65012

hypothetical protein b2382 - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 18-Sep-1998  
 C:Accession: C65012  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: C65012  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-285 <BLAT>  
 A:Cross-references: GB:AE000326; GB:U00096; NID:g1788718; PID:g1788725; UWGP:b2382  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Superfamily: hypothetical protein b2382

Query Match 68.8%; Score 33; DB 2; Length 285;  
 Best Local Similarity 55.6%; Pred. No. 28;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYVRV 9

|:|:|:|:|

Db 225 REIEYVRV 233

## RESULT 7

hypothetical protein HI0109 - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 10-Oct-1997  
 C:Accession: F64001  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630  
 A:Accession: F64001  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-113 <TIGR>  
 A:Cross-references: GB:U32696; GB:L42023; NID:g1573057; PID:g1573066; TIGR:HI0109

Query Match 68.8%; Score 33; DB 2; Length 113;  
 Best Local Similarity 62.5%; Pred. No. 11;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRV 9

|:|:|:|:|

Db 79 FVEIIRIV 86

## RESULT 8

hypothetical protein (internalin region) - Listeria monocytogenes  
 C:Species: Listeria monocytogenes  
 C:Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 18-Jun-1993  
 C:Accession: C39930  
 R:Gaillard, J.L.; Berche, P.; Frehel, C.; Goulin, E.; Cossart, P. Cell 65, 1127-1141, 1991  
 A:Title: Entry of Listeria monocytogenes into cells is mediated by internalin, a repeat  
 A:Reference number: A39930; MUID:91292517  
 A:Accession: C39930  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-630 <GAT>  
 A:Cross-references: GB:M67471

Query Match 68.8%; Score 33; DB 2; Length 630;  
 Best Local Similarity 71.4%; Pred. No. 62;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVRYVRV 7

|:|:|:|:|

Db 401 RYVRYVRV 407

## RESULT 9

ribosomal protein SL31 - Sulfolobus acidocaldarius  
 C:Species: Sulfolobus acidocaldarius  
 C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
 C:Accession: S53699; S41964  
 R:Moll, R.; Schmidtke, S.; Schaefer, G. Biochim. Biophys. Acta 1261, 315-318, 1995  
 A:Title: Nucleotide sequence of a gene cluster encoding ribosomal proteins in the the  
 A:Reference number: S53698; MUID:95226466  
 A:Accession: S53699  
 A:Molecule type: DNA  
 A:Residues: 1-129 <MOI>  
 A:Cross-references: EMBL:X77509; NID:g453437; PID:e300459; PID:g453439  
 C:Superfamily: rat ribosomal protein L31  
 C:Keywords: protein biosynthesis; ribosome

Query Match 66.7%; Score 32; DB 2; Length 129;  
 Best Local Similarity 55.6%; Pred. No. 19;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYVRV 9

|:|:|:|:|

Db 68 RAIKVRVY 76

## RESULT 10

white homolog - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 28-Aug-1998  
 C:Accession: G02068  
 R:Croop, J.M.; Tiller, G.; Fletcher, J.A.; Lux, M.; Raab, E.; Goldenson, D.; Arcinieg  
 Submitted to the EMBL Data Library, August 1995  
 A:Reference number: H00769  
 A:Accession: G02068  
 A:Status: Preliminary; translated from GB/EMBL/DBBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-638 <CRO>  
 A:Cross-references: EMBL:U34919; NID:g1314276; PID:g1314277  
 C:Genetics:  
 A:Gene: white  
 C:Superfamily: fruit fly white protein; ATP-binding cassette homology  
 C:Keywords: P-loop  
 F:61-253/Domain: ATP-binding cassette homology <ABC>  
 F:78-85/Region: nucleotide-binding motif A (P-loop)

Query Match 66.7%; Score 32; DB 2; Length 638;  
 Best Local Similarity 71.4%; Pred. No. 95;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRV 8

|:|:|:|:|

Db 558 YISYVRV 564

## RESULT 11

hypothetical protein BB0714 - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998

C:Accession: A70189  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt;  
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: A70189  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-322 <KLE>  
 A:Cross-references: GB:AE001171; GB:AE000783; NID:92688640; PID:g2688650; TIGR:BB0714  
 C:Experimental source: strain B31  
 C:Superfamily: hypothetical protein BB0714; tetratricopeptide repeat homology  
 F:122-155/Domain: tetratricopeptide repeat homology <TT1>  
 F:161-193/Domain: tetratricopeptide repeat homology #status atypical <TT2>  
 F:194-227/Domain: tetratricopeptide repeat homology <TT3>  
 F:240-273/Domain: tetratricopeptide repeat homology <TT4>

Query Match 66.7%; Score 32; DB 2; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 48;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVY 8  
 |||||  
 Db 4 RYLYRYFY 11

RESULT 12  
 S32879  
 lipA protein - Neisseria meningitidis  
 C:Species: Neisseria meningitidis  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 03-Aug-1995  
 C:Accession: S32879; S28077  
 R:Frosch, M.; Mueller, A.  
 Mol. Microbiol. 8, 483-493, 1993  
 A:Title: Phospholipid substitution of capsular polysaccharides and mechanisms of capsule  
 A:Reference number: S32879  
 A:Accession: S32879  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-400 <PRO>  
 A:Cross-references: EMBL:Z13995  
 C:Genetics:  
 A:Gene: lipA

Query Match 66.7%; Score 32; DB 2; Length 400;  
 Best Local Similarity 55.6%; Pred. No. 59;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVY 9  
 |||||  
 Db 134 RYLOYRNL 142

RESULT 13  
 H70679  
 probable membrane protein with some - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998  
 C:Accession: H70679  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 ; Connor, R.; Davies, R.; Devlin, K.; Beldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: H70679  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-481 <COL>  
 A:Cross-references: GB:Z81451; GB:AL123456; NID:g3261662; PID:e280499; PID:g1666150  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV2434c

Query Match 66.7%; Score 32; DB 2; Length 481;  
 Best Local Similarity 75.0%; Pred. No. 71;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVYRVY 8  
 |||||  
 Db 348 RYARLVRY 355

RESULT 14  
 C64347  
 hypothetical protein MJ0379 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 10-Oct-1997  
 C:Accession: C64347  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Raine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
 A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: C64347  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-208 <BUL>  
 A:Cross-references: GB:U67490; GB:L77117; NID:g1591074; PID:g1591084; TIGR:MJ0379; PI  
 C:Genetics:  
 A:Map position: REV345512-344886  
 A:Start codon: TTG

Query Match 66.7%; Score 32; DB 2; Length 208;  
 Best Local Similarity 71.4%; Pred. No. 31;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYRVYRVY 9  
 |||||  
 Db 2 RYRMRYI 8

RESULT 15  
 S26004  
 18S rRNA intron 1 protein - liverwort (Marchantia polymorpha) mitochondrion  
 C:Species: mitochondrion Marchantia polymorpha  
 C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 17-Mar-1999  
 C:Accession: S26004  
 R:Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.;  
 J. Mol. Biol. 223, 1-7, 1992  
 A:Title: Gene organization deduced from the complete sequence of liverwort *Marchantia*  
 A:Reference number: S25941; MUID:92114051  
 A:Accession: S26004  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-502 <ODA>  
 A:Cross-references: EMBL:M68929; NID:g786182; PID:g786243  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: mitochondrion

Query Match 66.7%; Score 32; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 3 VRYRY 8  
| | | | |  
Db 272 VRYRY 277

Search completed: September 7, 1999, 23:06:24  
Job time: 2472 sec

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		ID	Description
		Match	Length		
1	35	72.9	110	1	RK23_PORPU
2	34	70.8	675	1	KSC5_ECOLI
3	33	68.8	600	1	SPO8_YEAST
4	33	68.8	630	1	INLB_LISMO
5	33	68.8	113	1	Y109_HAEIN
6	33	68.8	1095	1	YOH5_YEAST
7	33	68.8	285	1	YPDC_ECOLI
8	32	66.7	507	1	INOL_CITPA
9	32	66.7	400	1	LIPA_NEIME
10	32	66.7	858	1	MISS_SCHPO
11	32	66.7	674	1	WHIT_HUMAN
12	32	66.7	666	1	WHIT_MOUSE
13	32	66.7	208	1	Y379_METJA
14	32	66.7	502	1	YMA0_MARPO
15	32	66.7	129	1	YRL1_SULAC
16	31	64.6	239	1	CC16_SCHPO
17	31	64.6	813	1	FGR2_XENLA
18	31	64.6	730	1	PL0D_CAEEL
19	30	62.5	20	1	AMP_FUSNU
20	30	62.5	1603	1	ARO1_EMENI
21	30	62.5	511	1	CIKD_HUMAN
22	30	62.5	511	1	CIKD_MOUSE
23	30	62.5	585	1	CIKD_RAT
24	30	62.5	570	1	HEMA_NDVA
25	30	62.5	571	1	HEMA_NDVH3
26	30	62.5	571	1	HEMA_NDVI
27	30	62.5	571	1	HEMA_NDVM
28	30	62.5	436	1	RBL_EUGST
29	30	62.5	163	1	TPX1_STRPN
30	30	62.5	1118	1	UPBR_HUMAN
31	30	62.5	340	1	YDDR_ECOLI
32	29	60.4	127	1	ATP2_BACP3
33	29	60.4	249	1	CAMT_EUCGU
34	29	60.4	241	1	CAMT_PETCR
35	29	60.4	247	1	CAMT_POPTM
36	29	60.4	242	1	CAMT_VITVI
37	29	60.4	1051	1	CARB_SULSO
38	29	60.4	1333	1	CC25_CANAL
39	29	60.4	351	1	COBJ_METTH
40	29	60.4	448	1	CR51_BACSH
41	29	60.4	448	1	CR53_BACSH
42	29	60.4	856	1	ENAF_FVNZT
43	29	60.4	445	1	ENAF_YEAST

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-K5;  
RC MEDLINE; 93388530.  
RX PAZZANI C., ROSENOW C., BOULNOIS G.J., BRONNER D., JANN K.  
RA ROBERTS I.S.;  
RT 'Molecular analysis of region 1 of the *Escherichia coli* K5  
RT gene cluster: a region encoding proteins involved in cell

RT expression of capsular polysaccharide.";

RL J. BACTERIOL. 175:5978-5983(1993).

RN [2]

RP SEQUENCE OF 1-8 FROM N.A.

RC STRAIN=K5;

RX MEDLINE; 95180691.

RA ROSENOW C., ROBERTS I.S., JANN K.;

RT "Isolation from recombinant Escherichia coli and characterization of

RT CMP-kdo synthetase, involved in the expression of the capsular K5

RT polysaccharide (K-CKS).";

RL FEMS MICROBIOL. LETT. 125:159-164(1995).

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).

CC -!- SIMILARITY: SOME, TO THE R.MELLORI LIPOPOLYSACCHARIDE PROCESSING

CC PROTEIN LP2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X74567; G397408; -

DR EMBL; S76943; E193729; -

KW POLYSACCHARIDE TRANSPORT; TRANSPORT

SQ SEQUENCE 675 AA; 75771 MW; 663ADIC3 CRC32;

Query Match 70.8%; Score 34; DB 1; Length 675;

Best Local Similarity 62.5%; Pred. No. 19;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRYVRYVY 9

DB 293 YLYRCYII 300

RESULT 3

SP08\_YEAST

ID SPO8\_YEAST STANDARD; PRT; 600 AA.

AC P41833;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE TRANSCRIPTIONAL REGULATOR SPO8.

GN SPO8 OR IME4 OR IGL192W OR G1337.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

CC SACCHAROMYCETACEAE; SACCHAROMYCES.

CC [1]

CC SEQUENCE FROM N.A.

RA SAKURAI M.;

RN SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.

CC [2]

CC SEQUENCE FROM N.A.

RC STRAIN=KKV5;

RA CLANCY M.J.;

RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC [3]

CC SEQUENCE FROM N.A.

RC STRAIN=S288C / FV1679;

RX MEDLINE; 97197971.

RA COGLIEVINA M., KLIMA R., BERTANI I., DELNERI D., ZACCARIA P.,

RA BRUSCHI C.V.;

RT "Sequencing of a 40.5 kb fragment located on the left arm of

RT chromosome VII from *Saccharomyces cerevisiae*.";

RL YEAST 13:55-64(1997).

CC -!- FUNCTION: POSITIVE TRANSCRIPTIONAL REGULATOR FOR IME2.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; D23721; G471168; -

DR EMBL; U30859; G945046; -

DR EMBL; X91837; E203623; -

DR EMBL; Z72714; E243795; -

SGD; L0000863; IME4

KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.

SQ SEQUENCE 600 AA; 69395 MW; B41A231B CRC32;

Query Match 70.8%; Score 34; DB 1; Length 600;

Best Local Similarity 44.4%; Pred. No. 17;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRYVRYV 9

DB 272 RYIHYLQVI 280

RESULT 4

INLB\_LISMO

ID INLB\_LISMO STANDARD; PRT; 630 AA.

AC P25147;

DT 01-MAY-1992 (REL. 22, CREATED)

DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)

DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)

DE INTERNALIN B PRECURSOR.

GN INLB.

OS LISTERIA MONOCYTOGENES.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;

CC LISTERIA.

CC [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91292517.

RA GAILLARD J.-L., BERGE P., FREHEL C., GOUIN E., COSSART P.;

RT "Entry of *L. monocytogenes* into cells is mediated by internalin, a

RT repeat protein reminiscent of surface antigens from gram-positive

RT cocci.";

RL CELL 65:1127-1141(1991).

CC -!- FUNCTION: MEDIATES THE ENTRY OF *LISTERIA MONOCYTOGENES* INTO CELLS.

CC -!- SIMILARITY: TO INTERNALIN A.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M67471; G149675; -

DR PIR; C39930; C39930.

DR PFAM; PF00560; LRR; 3.

DR REPEAT; SIGNAL.

ET SIGNAL

ET CHAIN 1 30

ET DOMAIN 31 630

ET REPEAT 63 238

ET REPEAT 85 106

ET REPEAT 107 128

ET REPEAT 129 150

ET REPEAT 151 172

ET REPEAT 173 194

ET REPEAT 195 216

ET REPEAT 217 238

ET DOMAIN 399 630

ET REPEAT 399 465

ET REPEAT 559 630

POTENTIAL.

INTERNALIN B.

8 X APPROXIMATE TANDEM REPEATS, TYPE A.

```

SQ SEQUENCE 630 AA; 71220 MW; CA2F6E58 CRC32;

Query Match 68.8%; Score 33; DB 1; Length 630;
Best Local Similarity 71.4%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVYR 7
Db 401 RYKVIK 407
||||:|

RESULT 5
YI09_HAEIN STANDARD; PRT; 113 AA.
AC P43943;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL PROTEIN HI0109 PRECURSOR.
GN HI0109.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA MCENNEY K., SUTTON G., FITZTHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODIE A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DE EMBL; U32696; G1573066; -.
DR TIGR; HI0109; -.
KW HYPOTHETICAL PROTEIN; SIGNAL.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 113 HYPOTHETICAL PROTEIN HI0109.
SQ SEQUENCE 113 AA; 13379 MW; 5321F58F CRC32;

Query Match 68.8%; Score 33; DB 1; Length 113;
Best Local Similarity 62.5%; Pred. No. 4.4;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRVYRVYV 9
Db 79 FVEYIRV 86
||||:|

RESULT 6
YOH5_YEAST STANDARD; PRT; 1095 AA.
ID YOH5_YEAST
AC Q08234;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DDG-GLK INTERGENIC REGION.
DE YPDC.
GN YPDC.

SQ SEQUENCE 630 AA; 71220 MW; CA2F6E58 CRC32;

DE GN YOLO75C.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;
OC SACHAROMYCETACEAE; SACHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97321807.
RA TZERMIAN M., KATSIOLOU C., ALEXANDRAKI D.;
RT "Sequence analysis of a 33.2 kb segment from the left arm of yeast
RT chromosome XV reveals eight known genes and ten new open reading
RT frames including homologues of ABC transporters, inositol
RT phosphatases and human expressed sequence tags.";
RL YEAST 13:583-589(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DE EMBL; Z74817; E251877; -.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_trap; 1.
KW HYPOTHETICAL PROTEIN; ATP-BINDING; TRANSMEMBRANE; GLYCOPROTEIN;
KW TRANSPORT.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 354 374 POTENTIAL.
FT TRANSMEM 406 426 POTENTIAL.
FT TRANSMEM 840 860 POTENTIAL.
FT TRANSMEM 871 891 POTENTIAL.
FT TRANSMEM 922 942 POTENTIAL.
FT TRANSMEM 949 969 POTENTIAL.
FT TRANSMEM 978 998 POTENTIAL.
FT TRANSMEM 1009 1029 POTENTIAL.
FT TRANSMEM 1068 1088 POTENTIAL.
FT NP_BIND 528 535 ATP (POTENTIAL).
FT CARBOHYD 142 142 POTENTIAL.
FT CARBOHYD 150 150 POTENTIAL.
FT CARBOHYD 172 172 POTENTIAL.
FT CARBOHYD 329 329 POTENTIAL.
FT CARBOHYD 784 784 POTENTIAL.
FT CARBOHYD 863 863 POTENTIAL.
SQ SEQUENCE 1095 AA; 123507 MW; 24679CD6 CRC32;

Query Match 68.8%; Score 33; DB 1; Length 1095;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRVYRVYV 9
Db 351 YVRWIKYI 358
||||:|

RESULT 7
YPDC_ECOLI STANDARD; PRT; 285 AA.
ID YPDC_ECOLI
AC P77396;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DDG-GLK INTERGENIC REGION.
DE YPDC.
GN YPDC.

```

```

OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RC MEDLINE: 97426617
RA BLATTNER F.R., FLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,
RA IKEMOTO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K., KASAI H.,
RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,
RA MASUDA S., MIKI T., MIZOBUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,
RA NASHIMOTO H., NISHIO Y., OSHIMA T., SAITO N., SAMPEI G., SEKI Y.,
RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE ARAC/XLYS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000326; G1788725; -.
CC EMBL; D90868; G1799793; -.
CC ECOGENE; EG14150; YPDC.
CC PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
CC PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
CC PFAM; PF00165; HTH_2; 1.
CC KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.
CC FT DNA_BIND 200 219 H-T-H MOTIF (BY SIMILARITY).
CC SEQUENCE 285 AA; 32355 MW; 61FFAA84 CRC32;

Query Match 68.8%; Score 33; DB 1; Length 285;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9
Db 225 RFIEYRVWV 233
|:::|::|:|

RESULT 8
ID INOL_CITPA STANDARD; PRT; 507 AA.
AC PA2802;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MYO-INOSITOL-1-PHOSPHATE SYNTHASE (EC 5.5.1.4) (IPS).
OS CITRUS PARADISI (GRAPFRUIT).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC SAPINDALES; RUTACEAE; CITRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF.
RX MEDLINE: 95148748.
RA ABU-ABIED M., HOLLAND D.;
RT "The gene c-inol from Citrus paradisi is highly homologous to turl
RT and inol from yeast and Spirodela encoding for myo-inositol phosphate

```

```

synthase.";
RL PLANT PHYSIOL. 106:1689-1689(1994).
CC -!- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = 1L-MYO-INOSITOL
CC 1-PHOSPHATE.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: INOSITOL BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE MYO-INOSITOL-1-PHOSPHATE SYNTHASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z32632; G602565; -.
CC KW PHOSPHOLIPID BIOSYNTHESIS; INOSITOL BIOSYNTHESIS; ISOMERASE; NAD.
CC SEQUENCE 507 AA; 56334 MW; 60127CC0 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 507;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9
Db 382 RWIKYVPYV 390
|:::|::|:|

RESULT 9
ID LIPA_NEIME STANDARD; PRT; 400 AA.
AC Q05013;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE CAPSULE POLYSACCHARIDE MODIFICATION PROTEIN LIPA.
GN LIPA.
OS NEISSERIA MENINGITIDIS.
OC BACTERIA; PROTEOBACTERIA; BETA SUBDIVISION; NEISSERIACEAE; NEISSERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE B1940 / SEROGROUP B;
RX MEDLINE: 93316845.
RA PROSCH M., MUELLER A.;
RT "Phospholipid substitution of capsular polysaccharides and mechanisms
RL of capsule formation in Neisseria meningitidis.";
RL MOL. MICROBIOL. 8:483-493(1993).
CC -!- FUNCTION: INVOLVED IN THE PHOSPHOLIPID MODIFICATION OF THE
CC CAPSULAR POLYSACCHARIDE, A STRONG REQUIREMENT FOR ITS
CC TRANSLOCATION TO THE CELL SURFACE.
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (CYTOPLASMIC
CC SIDE) (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z13995; -. NOT_ANNOTATED_CDS.
CC FIR; S28077; S28077.
CC PIR; S32879; S32879.
CC KW INNER MEMBRANE; POLYSACCHARIDE TRANSPORT; TRANSPORT.
CC SEQUENCE 400 AA; 45106 MW; B99FB701 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 400;

```

```

Best Local Similarity 55.6%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 9
   ||:| | |
Db 134 RYQYYSRYL 142

RESULT 10
MISS_SCHPO
ID MISS_SCHPO STANDARD; PRT; 868 AA.
AC P49731;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MISS PROTEIN.
GN MISS.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95170112.
RA TAKAHASHI K., YAMADA H., YANAGIDA M.;
RT "Fission yeast minichromosome loss mutants mis cause lethal
   aneuploidy and replication abnormality.";
RL MOL. BIOL. CELL 5:1145-1158(1994).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D31960; G829086; -
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS50051; MCM_2; 1.
DR PFAM; PF00493; MCM; 1.
KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW DNA REPLICATION; CELL CYCLE; ATP-BINDING.
FT DOMAIN 426 633 MCM.
FT NP_BIND 476 483 ATP (POTENTIAL).
FT CONFLICT 476 483
SQ SEQUENCE 868 AA; 96753 MW; EAF553B0 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 868;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVYV 7
   ||:| | |
Db 654 RYRYVAR 660

RESULT 11
WHIT_HUMAN
ID WHIT_HUMAN STANDARD; PRT; 674 AA.
AC P45844;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE WHITE PROTEIN HOMOLOG (ATP-BINDING CASSETTE TRANSPORTER 8).
GN ABC8 OR WH1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.

```

```

TISSUE-RETINA;
RC MEDLINE; 96256850.
RA CHEN H.M., ROSSIER C., LALIOI M.D., LYNN A., CHAKRAVARTI A.,
RA PERRIN G., ANTONARAKIS S.E.;
RT "Cloning of the cDNA for a human homologue of the Drosophila white
   gene and mapping to chromosome 21q22.3.";
RL AM. J. HUM. GENET. 59:66-75(1996).
RN [2]
RP SEQUENCE OF 37-674 FROM N.A.
RC TISSUE-FETAL BRAIN;
RX MEDLINE; 97186700.
RA CROOP J.M., TILLER G., FLETCHER J.A., LUX M., RAAB E., GOLDENSON D.,
RA ARCINIEGAS S., SON D., WU R.;
RT "Isolation and characterization of a mammalian homolog of the
   Drosophila white gene.";
RL GENE 185:77-85(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SEVERAL TISSUES.
CC -!- ALTERNATIVE PRODUCTS: IT IS POSSIBLE THAT TWO DIFFERENT FORMS THAT
   DIFFER IN THE PRESENCE OR ABSENCE OF A 12 RESIDUE REGION ARE
   PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
   (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X91249; E218444; -
DR EMBL; U34919; G1314277; -
DR MIM; 603076; -
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
KW ATP-BINDING; TRANSMEMBRANE; TRANSPORT; ALTERNATIVE SPLICING.
FT NP_BIND 114 121 ATP (POTENTIAL).
FT TRANSMEM 423 441 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 503 521 POTENTIAL.
FT TRANSMEM 530 551 POTENTIAL.
FT TRANSMEM 564 582 POTENTIAL.
FT TRANSMEM 646 665 POTENTIAL.
FT VARSPIC 371 382 MISSING (IN A SHORTER FORM).
FT CONFLICT 444 444 T -> A (IN REF. 2).
FT CONFLICT 529 529 R -> A (IN REF. 2).
FT CONFLICT 664 664 L -> F (IN REF. 2).
SQ SEQUENCE 674 AA; 75169 MW; 536156C1 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 674;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRVYRVYV 8
   |::| | | |
Db 594 YISYVRY 600

RESULT 12
WHIT_MOUSE
ID WHIT_MOUSE STANDARD; PRT; 666 AA.
AC Q64343;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE WHITE PROTEIN HOMOLOG (ATP-BINDING CASSETTE TRANSPORTER 8).
GN ABC8 OR WH1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

```

```

RA SEQUENCE FROM N.A.
RX MEDLINE: 97186700.
RA CROOP J.M., TILLER G.E., FLETCHER J.A., LUX M.L., RAAB E.,
RA GOLDENSON D., SON D., ARCINIEGAS S., WU R.;
RT "Isolation and characterization of a mammalian homolog of the
RT Drosophila white gene."
RL GENE 185:77-85(1997).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=DBA/2;
RX MEDLINE: 96359154.
RA SAVARY S., DENIZOT F., LUCIANI M.-F., MATTEI M.-G., CHIMINI G.;
RT "Molecular cloning of a mammalian ABC transporter homologous to
RT Drosophila white gene."
RL MAMM. GENOME 7:673-676(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN, THYMUS, LUNG,
CC ADRENALS, SPLEEN AND PLACENTA. LITTLE OR NO EXPRESSION IN LIVER,
CC KIDNEY, HEART, MUSCLE OR TESTES.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U34920; G1314279; -
DR EMBL: Z48745; E156982; -
DR MGD; MGI:107704; ABC8.
DR PROSITE; PS00211; ABC TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
DR HSP; P30533; INRE.
KW ATP-BINDING; TRANSMEMBRANE; TRANSPORT.
FT NP_BIND 118 125 ATP (POTENTIAL).
FT TRANSMEM 415 433 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 495 513 POTENTIAL.
FT TRANSMEM 522 543 POTENTIAL.
FT TRANSMEM 556 574 POTENTIAL.
FT TRANSMEM 638 657 POTENTIAL.
SQ SEQUENCE 666 AA; 74033 MW; AB4B792F CRC32;

Query Match 66.7%; Score 32; DB 1; Length 666;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYRVY 8
Db 586 YISYRVY 592
I: |||||

RESULT 13
Y379_METJA
ID Y379_METJA STANDARD; PRT; 208 AA.
AC Q57824;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL PROTEIN MJ0379.
GN MJ0379.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
CC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.

```

```

RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER K., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOSESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL SCIENCE 273:1058-1073(1996).
CC -1- SIMILARITY: IN THE C-TERMINAL, TO M.JANNASCHII MJ0723.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67490; G1591084; -
DR TIGR; MJ0379; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 208 AA; 24080 MW; 5DD82FAD CRC32;

Query Match 66.7%; Score 32; DB 1; Length 208;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRYRVYV 9
Db 2 VRYMYI 8
I: |||:|

RESULT 14
YM40_MARPO
ID YM40_MARPO STANDARD; PRT; 502 AA.
AC P38478;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 57.7 KD PROTEIN IN 18S RNA INTRON 1 (ORF 502).
GN YMF40.
OS MARCHANTIA POLYMORPHA (LIVERWORT).
OG MITOCHONDRION.
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; MARCHANTIOPSIDA;
OC MARCHANTIALES; MARCHANTIACEAE; MARCHANTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92114051.
RA ODA K., YAMATO K., OHTA E., NAKAMURA Y., TAKEMURA M., NOZATO N.,
RA AKASHI K., KANEGAE T., OGURA Y., KOHCHI T., OHYAMA K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome."
RL J. MOL. BIOL. 223:1-7(1992).
CC -1- SIMILARITY: TO GROUP II INTRON MATURASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M68929; G786243; -
DR PIR; S26004; S26004.
DR PFAM; PF00078; rvt; 1.
DR MENDEL; 2114; MARPO:ymf40.1.
KW MITOCHONDRION; HYPOTHETICAL PROTEIN.

```



SQ SEQUENCE 502 AA; 57722 MW; A8621A83 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 502;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRYVRY 8  
|:|:|:|:|  
Db 272 VRYVRY 277

RESULT 15

YRLI\_SULAC STANDARD; PRT; 129 AA.  
AC P38618;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 15.2 KD PROTEIN IN RL46-RLX INTERGENIC REGION (A01).  
OS SULFOLOBUS ACIDOCALDARIUS.  
OC ARCHAEA; CRENARCHAEOTA; SULFOLOBALES; SULFOLOBUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 639;  
RX MEDLINE; 95226466.  
RA MOLL R., SCHAEFER G., SCHMIDTKE S.;  
RT "Nucleotide sequence of a gene cluster encoding ribosomal proteins in  
the thermophilic crenarchaeon Sulfolobus acidocaldarius";  
RL BIOCHIM. BIOPHYS. ACTA 1261:315-318(1995).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X77509; G453439; -  
DR EMBL; X77509; E300459; -  
DR PIR; S41964; S41964.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 129 AA; 15196 MW; E346121E CRC32;

Query Match 66.7%; Score 32; DB 1; Length 129;  
Best Local Similarity 55.6%; Pred. No. 7.9;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVRVRYV 9  
|:|:|:|:|  
Db 68 RAIKYVRYM 76

Search completed: September 7, 1999, 23:50:12  
Job time: 1951 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:38 ; Search time 148.39 Seconds  
(without alignments)  
3.733 Million cell updates/sec

Title: US-09-124-280A-9  
Perfect score: 48  
Sequence: 1 RYRVYRVV 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTRMBL10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	72.9	426	11 055082	055082 mus musculus
2	34	70.8	261	1 026968	026968 methanobact
3	34	70.8	394	2 085335	085335 escherichia
4	34	70.8	438	4 075783	075783 homo sapien
5	34	70.8	149	13 Q9YHK6	Q9YHK6 ginglymosto
6	33	68.8	415	2 Q92739	Q92739 chlamydia p
7	33	68.8	533	4 015317	015317 homo sapien
8	33	68.8	337	5 017433	017433 caenorhabdi
9	33	68.8	1928	5 097291	097291 plasmodium
10	33	68.8	325	12 011441	011441 urochloa ho
11	32	66.7	322	2 051656	051656 borrelia bu
12	32	66.7	481	2 P71915	P71915 mycobacteri
13	32	66.7	306	3 074870	074870 schizosacch
14	32	66.7	900	3 074925	074925 schizosacch
15	32	66.7	232	4 043576	043576 homo sapien
16	32	66.7	200	5 020126	020126 caenorhabdi
17	32	66.7	425	5 018923	018923 caenorhabdi
18	32	66.7	296	5 093456	093456 caenorhabdi
19	32	66.7	165	5 096406	096406 plasmodium
20	32	66.7	785	8 033759	033759 allomyces m
21	32	66.7	636	8 099479	099479 pavlova lut
22	32	66.7	415	10 081088	081088 tradescantl
23	32	66.7	857	12 066956	066956 feline immu
24	32	66.7	852	12 066959	066959 feline immu
25	31	64.6	223	1 058882	058882 pyrococcus
26	31	64.6	645	2 045773	045773 bacteroides
27	31	64.6	170	2 P76492	P76492 escherichia
28	31	64.6	301	2 007156	007156 mycobacteri
29	31	64.6	514	2 067596	067596 aquifex aeo

30	31	64.6	346	2 069652	069652 mycobacteri
31	31	64.6	1166	2 092757	092757 chlamydia p
32	31	64.6	829	5 018253	018253 caenorhabdi
33	31	64.6	414	5 045571	045571 caenorhabdi
34	31	64.6	1259	8 035058	035058 marchantia
35	31	64.6	336	10 039218	039218 arabidopsis
36	31	64.6	391	10 042532	042532 arabidopsis
37	31	64.6	396	12 056869	056869 gallid herp
38	30	62.5	228	2 034903	034903 bacillus su
39	30	62.5	249	2 083299	083299 treponema p
40	30	62.5	280	2 085861	085861 sphingomona
41	30	62.5	37	2 049100	049100 mycoplasma
42	30	62.5	238	2 045591	045591 bacillus su
43	30	62.5	862	5 022354	022354 caenorhabdi
44	30	62.5	477	5 019988	019988 caenorhabdi
45	30	62.5	1033	5 044323	044323 forficula a

ALIGNMENTS

RESULT 1  
O55082  
ID O55082 PRELIMINARY; PRT: 426 AA.  
AC O55082;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 20 (EC 3.1.3.48)  
DE (PHOSPHOTYROSINE PHOSPHATASE) (PTPASE).  
GN PTPN20.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TESTIS;  
RX MEDLINE; 98070510.  
RA OHSUGI M., KURAMOTOCHI S., MATSUDA S., YAMAMOTO T.;  
RT "Molecular cloning and characterization of a novel cytoplasmic  
RT protein-tyrosine phosphatase that is specifically expressed in  
RT spermatocytes.";  
RL J. Biol. Chem. 272:33092-33099(1997).  
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN  
CC TYROSINE + PHOSPHATE.  
DR EMBL; D64141; BAA23761.1; .  
DR MGD; MGI:1196295; PTPN20.  
DR PFM; PFM0102; Y\_phosphatase; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
KW Hydrolase.  
SQ SEQUENCE 426 AA; 49118 MW; 58F1AA05 CRC32;

Query Match 72.9%; Score 35; DB 11; Length 426;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 YRVYRVV 9  
DB 339 FIKYRVV 346

RESULT 2  
O26968  
ID O26968 PRELIMINARY; PRT: 261 AA.  
AC O26968;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)  
DE CONSERVED PROTEIN.  
GN MTH882.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;

```

OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE: 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDRIDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL: AF000864; AAB85380.1; -.
SQ SEQUENCE 261 AA; 26929 MW; 53F0B865 CRC32;

Query Match 70.8%; Score 34; DB 1; Length 261;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRYVRYV 9
Db 241 VRYIRYI 247

RESULT 3
O85335 SEQUENCE FROM N.A.
ID O85335 PRELIMINARY; PRT; 394 AA.
AC O85335;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE O ANTIGEN POLYMERASE WZY.
GN WZY.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C664-1992;
RX MEDLINE: 98339851.
RA WANG L., REEVES P.R.;
RT "Organization of Escherichia coli O157 O antigen gene cluster and
RT identification of its specific genes.";
RL Infect. Immun. 66:3545-3551(1998).
DR EMBL: AF061251; AAC32340.1; -.
SQ SEQUENCE 394 AA; 45647 MW; EF0C67FB CRC32;

Query Match 70.8%; Score 34; DB 2; Length 394;
Best Local Similarity 57.1%; Pred. No. 65;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRYVRYV 8
Db 117 YRYIKY 123

RESULT 4
O75783 SEQUENCE FROM N.A.
ID O75783 PRELIMINARY; PRT; 438 AA.
AC O75783;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE RHOMBOID-RELATED PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE: 98324821.
RA PASCALL J.C., BROWN K.D.;
RT "Characterization of a mammalian cDNA encoding a protein with high
RT sequence similarity to the Drosophila regulatory protein Rhomboid.";
RL FEBS Lett. 429:337-340(1998).
DR EMBL: Y17108; CAA76629.1; -.
SQ SEQUENCE 438 AA; 48314 MW; E05F94D5 CRC32;

Query Match 70.8%; Score 34; DB 4; Length 438;
Best Local Similarity 75.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVRYVY 8
Db 168 RYRVRYVAY 175

RESULT 5
O9YHK6 PRELIMINARY; PRT; 149 AA.
ID O9YHK6;
AC O9YHK6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE ANTIGEN RECEPTOR (FRAGMENT).
GN NAR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Orectolobiformes; Ginglymostomatidae; Ginglymostoma.
RN [1]
RP SEQUENCE FROM N.A.
RA DIAZ M., GREENBERG A.S., FLAJNIK M.F.;
RT "Somatic Hypermutation of the New Antigen Receptor Gene (NAR) in the
RT Nurse Shark Does Not Generate the Repertoire: Possible Role in
RT Antigen-Driven Reactions in the Absence of Germinal Centers.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14343-14348(1998).
DR EMBL: AF096029; AAC83731.1; -.
KW Receptor.
FT NON_TER 1
FT NON_TER 149
SQ SEQUENCE 149 AA; 16362 MW; 8F1AD358 CRC32;

Query Match 70.8%; Score 34; DB 13; Length 149;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVRYVY 9
Db 53 RYVETRYVY 61

RESULT 6
O9Z739 PRELIMINARY; PRT; 415 AA.
ID O9Z739;
AC O9Z739;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE ROD SHAPE PROTEIN.
GN RODA.
OS Chlamydia pneumoniae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiaophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RA KALMAN S., MITCHELL W., MARATHE R., LAMMEL C., FAN J., OLINGER L.,
RA GRIMWOOD J., DAVIS R.W., STEPHENS R.S.;
RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001667; AAD19005.1; -.

```

```

SQ SEQUENCE 415 AA; 46662 MW; 3A4ABBC4 CRC32;

Query Match
Best Local Similarity 68.8%; Score 33; DB 2; Length 415;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVRYRYV 9
DB 38 RYHKYRYV 46

RESULT 7
ID O15317 PRELIMINARY; PRT; 533 AA.
AC O15317;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE RNA POLYMERASE III SUBUNIT.
GN RPCG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97315201.
RT WANG Z., ROEDER R.G.;
RT "Three human RNA polymerase III-specific subunits form a subcomplex
RT with a selective function in specific transcription initiation.";
RL Genes Dev. 11:1315-1326(1997).
DR EMBL; U93867; AAB63675.1; -.
SQ SEQUENCE 533 AA; 60629 MW; 851C71CF CRC32;

Query Match
Best Local Similarity 68.8%; Score 33; DB 4; Length 533;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVRYRYV 9
DB 84 RYRVRYRYI 92

RESULT 8
ID Q17433 PRELIMINARY; PRT; 337 AA.
AC Q17433;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 40.0 KD PROTEIN B0035.2 IN CHROMOSOME IV PRECURSOR.
GN B0035.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WHITE S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS A DNAJ-LIKE DOMAIN.
DR EMBL; Z73102; CAA97409.1; -.
DR WORMPEP; B0035.2; CE05161.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PFAM; PF00226; DnaJ; 1.
DR KW Hypothetical protein; Signal; Chaperone.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 337 HYPOTHETICAL PROTEIN B0035.2.
FT DOMAIN 34 107 DNAJ-LIKE.
SQ SEQUENCE 337 AA; 39992 MW; 5A8733EC CRC32;

SQ SEQUENCE 415 AA; 46662 MW; 3A4ABBC4 CRC32;

Query Match
Best Local Similarity 68.8%; Score 33; DB 5; Length 337;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVRYRYV 9
DB 317 RYKRYKRYM 325

RESULT 9
ID O97291 PRELIMINARY; PRT; 1928 AA.
AC O97291;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE MAL3P7.21 PROTEIN.
GN MAL3P7.21.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034559; CAB39037.1; -.
SQ SEQUENCE 1928 AA; 230190 MW; A8A8DEB CRC32;

Query Match
Best Local Similarity 68.8%; Score 33; DB 5; Length 1928;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRVRYRYV 9
DB 964 YVHYVHYV 971

RESULT 10
ID O11441 PRELIMINARY; PRT; 325 AA.
AC O11441;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE RNA REPLICASE (FRAGMENT).
GN PC1.
OS Urochloa hoja blanca virus.
OC Viruses; ssRNA negative-strand viruses; Tenuivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COSTA RICA;
RA DE MIRANDA J.R., MUNOZ M., CABEZAS E., HERNANDEZ M., ROOSSINCK M.J.,
RA WU R., ESPINOZA A.M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U82448; AAB58306.1; -.
FT NON_TER 1
SQ SEQUENCE 325 AA; 37992 MW; 80D2678A CRC32;

Query Match
Best Local Similarity 68.8%; Score 33; DB 12; Length 325;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YRVRYRYV 9
DB 260 VKYRYV 266

RESULT 11
ID O51656 PRELIMINARY; PRT; 322 AA.
AC O51656;
```

DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 38.6 KD PROTEIN.  
 GN B80714.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATHEY L., McDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE001171; AAC67064.1; .  
 DR TIGR; B80714; .  
 KW Hypothetical protein.  
 SQ SEQUENCE 322 AA; 38625 MW; 79D82CBA CRC32;

Query Match 66.7%; Score 32; DB 2; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRY 8  
 |||||  
 DB 4 RYLYRYFY 11

RESULT 12  
 ID P71915 PRELIMINARY; PRT; 481 AA.  
 AC P71915;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 52.3 KD PROTEIN.  
 GN MTCY428.12.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA DEVLIN K., CHURCHER C.M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE; 96181548.  
 RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,  
 RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,  
 RA COLE S.T.;  
 RT "An integrated map of the genome of the tubercle bacillus, Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium leprae."  
 RT Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).  
 RL EMBL; Z81451; CAB03785.1; .  
 DR PFAM; PF00924; UPF0003; 1.  
 KW Hypothetical protein.

SQ SEQUENCE 481 AA; 52272 MW; F54759A9 CRC32;  
 Query Match 66.7%; Score 32; DB 2; Length 481;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVYRY 8  
 |||||  
 DB 348 RYARLVRY 355

RESULT 13  
 ID 074870 PRELIMINARY; PRT; 306 AA.  
 AC 074870;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN SPC31H12.02C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA WOOD V., RAJANDREAM M.A., BARRELL B.G., RIEGER M.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031824; CAA21219.1; .  
 DR PFAM; PF01036; Bac\_rhodopsin; 1.  
 SQ SEQUENCE 306 AA; 34663 MW; 16578FD7 CRC32;

Query Match 66.7%; Score 32; DB 3; Length 306;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVYRY 9  
 |||||  
 DB 96 RMYVYRYI 104

RESULT 14  
 ID 074925 PRELIMINARY; PRT; 900 AA.  
 AC 074925;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE PUTATIVE VACUOLAR MEMBRANE PROTEIN.  
 GN SPC790.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., HILBERT H., DUESTERHOEFT A.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031855; CAA21292.1; .  
 SQ SEQUENCE 900 AA; 103719 MW; F96A69E7 CRC32;

Query Match 66.7%; Score 32; DB 3; Length 900;  
 Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYRVYRY 9  
 |||||  
 DB 629 IRYLRVY 635

## RESULT 15

O43576 PRELIMINARY: PRT; 232 AA.  
 AC O43576;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE WHITE PROTEIN HOMOLOG (FRAGMENT).  
 GN WHITE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE: 96207227.  
 RA ANDERSSON B., WENTLAND M.A., RICAPRENTE J.Y., LIU W., GIBBS R.A.;  
 RT "A double adaptor method for improved shotgun library  
 construction";  
 RL Anal. Biochem. 236:107-113(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE: 97264341.  
 RA YU W., ANDERSSON B., WORLEY K.C., MUZNY D.M., DING Y., LIU W.,  
 RA RICAPRENTE J.Y., WENTLAND M.A., LENNON G., GIBBS R.A.;  
 RT "Large-scale concatenation cDNA sequencing";  
 RL Genome Res. 7:353-358(1997).  
 DR EMBL; AF038175; AAB97364.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 232 AA; 26479 MW; 9E06238B CRC32;

Query Match 66.7%; Score 32; DB 4; Length 232;  
 Best Local Similarity 71.4%; Pred. NO. 85;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 YRYVRY 8  
 I: |||||  
 Db 152 YISYVRY 158

Search completed: September 7, 1999, 20:34:40  
 Job time: 19736 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:40 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-10

Perfect score: 56

Sequence: 1 KFFKFFKFFK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL\_10:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	71.4	980	5	017244
2	39	69.6	2010	5	096171
3	38	67.9	303	2	044421
4	38	67.9	1307	3	006412
5	38	67.9	395	5	017129
6	38	67.9	261	5	044194
7	38	67.9	738	5	097037
8	38	67.9	873	8	037370
9	37	66.1	192	2	051293
10	37	66.1	169	5	091319
11	37	66.1	592	5	017265
12	37	66.1	548	5	044959
13	37	66.1	619	5	018207
14	36	64.3	148	2	050972
15	36	64.3	409	2	072840
16	36	64.3	414	2	092N10
17	36	64.3	709	3	059824
18	36	64.3	430	5	001449
19	36	64.3	2206	5	096205
20	36	64.3	1014	5	096276
21	36	64.3	1827	5	097275
22	36	64.3	411	10	081438
23	36	64.3	89	12	039VI3
24	35	62.5	240	2	025697
25	35	62.5	240	2	092M55
26	35	62.5	615	3	087002
27	35	62.5	345	5	091014
28	35	62.5	261	5	076614
29	35	62.5	339	5	091891

30	35	62.5	552	8	092YM7
31	35	62.5	772	10	Q9ZPH7
32	34.5	61.6	448	2	033597
33	34.5	61.6	440	2	033616
34	34.5	61.6	441	2	033652
35	34.5	61.6	440	2	033671
36	34	60.7	427	2	033751
37	34	60.7	825	2	039319
38	34	60.7	649	2	084307
39	34	60.7	117	3	008205
40	34	60.7	328	5	021396
41	34	60.7	208	5	018525
42	34	60.7	807	5	017379
43	34	60.7	468	5	044439
44	34	60.7	4981	5	077372
45	34	60.7	495	5	019453

ALIGNMENTS

RESULT 1

ID	017244	PRELIMINARY;	PRT;	980 AA.
AC	017244;			
DT	01-JAN-1998 (TREMBlrel. 05, Created)			
DT	01-JAN-1998 (TREMBlrel. 05, Last sequence update)			
DT	01-NOV-1998 (TREMBlrel. 08, Last annotation update)			
DE	F53C2.6 PROTEIN.			
GN	F53C2.6.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;			
OC	Rhabditina; Rhabditoides; Rhabditidae; Peloderinae; Caenorhabditis.			
RC	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RX	MEDLINE: 94150718;			
RA	WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,			
RA	BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,			
RA	CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,			
RA	GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,			
RA	JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,			
RA	LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,			
RA	PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,			
RA	SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,			
RA	THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,			
RA	WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.; III of C.			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.			
RL	elegans.";			
RN	Nature 368:32-38(1994).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RA	BECKER M., WOHLDMANN P., BIEWALD T.;			
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RA	WATERSTON R.;			
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF025464; AAB71016.1;			
SQ	SEQUENCE 980 AA; 111010 MW; 4A69D72C CRC32;			

Query Match 71.4%; Score 40; DB 5; Length 980;  
Best Local Similarity 87.5%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EFKFKFF 9

DB 45 FFFKFF 52

```

RESULT 2
O96171 ID O96171 PRELIMINARY; PRT; 2010 AA.
AC O96171;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)
DE PHOSPHATASE (ACID PHOSPHATASE FAMILY).
GN PF0380C
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA GARDNER M.J., TATELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Science 282:1126-1132(1998).
DR EMBL: AE001391; AAC71865.1; -.
SQ SEQUENCE 2010 AA; 241791 MW; 46C82357 CRC32;

Query Match 69.6%; Score 39; DB 5; Length 2010;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKFKFKF 8
Db 801 KFKFKFKF 808

RESULT 3
Q44421 ID Q44421 PRELIMINARY; PRT; 303 AA.
AC Q44421;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE SUGAR-BINDING TRANSPORT PROTEIN.
OS Anaerococcus thermophilum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Anaerococcus group; Anaerococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z-1320;
RA ZVERLOV V., ASCHERL G., VELIKODVORSKAYA G., BRONNENMEIER K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z77855; CAB01403.1; -.
DR PFAM: PF00528; BPD_transp; 1.
SQ SEQUENCE 303 AA; 35132 MW; CF756347 CRC32;

Query Match 67.9%; Score 38; DB 2; Length 303;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFKFKFKF 9
Db 114 KFKFKFKF 122

RESULT 4
Q06412 ID Q06412 PRELIMINARY; PRT; 1307 AA.
AC Q06412;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE CHROMOSOME XII COSMID 9576.
GN L9576.5.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE: 97313267.
RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,
RA BENES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOEF A.,
RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEBLING U., HEUMANN K.,
RA HEUS-NEITZEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,
RA LOUIS E.J., MESSENGUY F., MEWES H.W., MOSGA T., MOSTL D.,
RA MULLER-AUER S., NENTWICH U., OBERWAIR B., PIRAVANDI E., POHL T.M.,
RA PORTELE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,
RA SCHARF M., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,
RA UNDERWOOD A.P., URRESTARAZU L.A., VANDENBOL M., VERHASSELT P.,
RA VIENENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUUT R., WEDLER E.,
RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOHEISEL J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XII."
RL Nature 387:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA FAVELLO A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WATERSTON R.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA CHERRY J.M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U20939; AAB67506.1; -.
DR PFAM: PF00780; CNH; 1.
SQ SEQUENCE 1307 AA; 149111 MW; AA414098 CRC32;

Query Match 67.9%; Score 38; DB 3; Length 1307;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKFKFKF 10
Db 1075 KFKFKFKF 1084

RESULT 5
O17129 ID O17129 PRELIMINARY; PRT; 395 AA.
AC O17129;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE F31F4.11 PROTEIN.
GN F31F4.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BEKKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

```

RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans."; Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BLANCHARD M., KRAMER J., ELLIOTT G., TWYMAN B.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF024503; AAB70387.1; -.  
 SQ SEQUENCE 395 AA; 46055 MW; 48E33092 CRC32;

Query Match 67.9%; Score 38; DB 5; Length 395;  
 Best Local Similarity 66.7%; Pred. No. 46;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFFK 10  
 I: | | | | |  
 Db 38 FWSFYKFFK 46

RESULT 6  
 ID O44194 PRELIMINARY; PRT; 261 AA.  
 AC O44194;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE T11F8.1 PROTEIN.  
 GN T11F8.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JERRE L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWKNEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans."; Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA JOHNSON D.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF036703; AAB88555.1; -.  
 SQ SEQUENCE 261 AA; 30914 MW; 7DBE6570 CRC32;

Query Match 67.9%; Score 38; DB 5; Length 261;

Best Local Similarity 66.7%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 FFKFFKFFK 10  
 I: | | | | |  
 Db 244 FYKFIQFFK 252

RESULT 7  
 ID O97037 PRELIMINARY; PRT; 738 AA.  
 AC O97037;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE PLC-DELTAH.  
 OS Hydra magnipapillata (Hydra).  
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;  
 OC Hydridae; Hydra.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=105;  
 RA KOYANAGI M., ONO K., SUGA H., IWABE N., MIYATA T.;  
 RT "Phospholipase C cDNAs from sponge and hydra: Antiquity of genes  
 involved in the inositol phospholipid signaling pathway.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB017513; BAA76278.1; -.  
 SQ SEQUENCE 738 AA; 84469 MW; DC8BCB99 CRC32;

Query Match 67.9%; Score 38; DB 5; Length 738;  
 Best Local Similarity 70.0%; Pred. No. 80;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
 I: | | | | | | |  
 Db 206 EFFKFFSFDK 215

RESULT 8  
 ID Q37370 PRELIMINARY; PRT; 873 AA.  
 AC Q37370;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE II (EC 1.9.3.1).  
 GN COXI/2.  
 OS Acanthamoeba castellanii (Amoeba).  
 OC Mitochondrion.  
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEFF (ATCC 30010);  
 RX MEDLINE; 95147275.  
 RA BURGER G., PLANTE I., LONERGAN K.M., GRAY M.W.;  
 RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba  
 castellanii: complete sequence, gene content and genome  
 organization."; J. Mol. Biol. 245:522-537(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEFF (ATCC 30010);  
 RA BURGER G.;  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O + 4  
 CC -1- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME  
 CC FERROCYTOCHROME C.  
 CC COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED  
 CC VIA HEME A AND CU(A) TO THE BINUCLEAR CENTER FORMED BY HEME A3  
 CC AND CU(B). SUBUNIT II BINDS CU(A) AND CYTOCHROME C.  
 CC -1- COFACTOR: COPPER A AND HEME GROUP.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. (MITOCHONDRIAL).

CC -1- SIMILARITY: TO OTHER MITOCHONDRIAL OR BACTERIAL COX2 SUBUNITS.  
 CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 DR EMBL; U12386; AAD11820.1; -;  
 DR PFAM; PF00115; COX1; 1.  
 DR PFAM; PF00116; COX2; 3.  
 DR PROSITE; PS00077; COX1; 1.  
 DR PROSITE; PS00078; COX2; 1.  
 KW Mitochondrion; Oxidoreductase; Copper; Transmembrane; Heme;  
 Respiratory chain.  
 FT METAL 803 803 COPPER A (PROBABLE).  
 FT METAL 807 807 COPPER A (PROBABLE).  
 FT METAL 836 836 COPPER A (PROBABLE).  
 FT METAL 840 840 COPPER A (PROBABLE).  
 FT METAL 844 844 COPPER A (PROBABLE).  
 FT METAL 847 847 COPPER A (PROBABLE).  
 FT METAL 847 847 COPPER A (PROBABLE).  
 SQ SEQUENCE 873 AA; 99213 MW; 654D0DD9 CRC32;  
 Query Match 67.9%; Score 38; DB 8; Length 873;  
 Best Local Similarity 87.5%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 FFKFFKFF 9  
 |||||  
 Db 530 FFKFSKFF 537  
 RESULT 9  
 ID 051293 PRELIMINARY; PRT; 192 AA.  
 AC 051293;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE CELL DIVISION PROTEIN (FTSJ).  
 GN BB0313.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE001138; AAC66697.1; -;  
 DR TIGR; BB0313; -;  
 KW Cell division.  
 SQ SEQUENCE 192 AA; 22167 MW; 1BF88CA4 CRC32;  
 Query Match 66.1%; Score 37; DB 2; Length 192;  
 Best Local Similarity 70.0%; Pred. No. 35;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKFFKFFK 10  
 |||  
 Db 160 KFKFYKFFK 169  
 RESULT 10  
 ID P91319 PRELIMINARY; PRT; 169 AA.  
 AC P91319;

DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE COSMID F53E10.  
 GN F53E10.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA BECK C., WAMSLEY P.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U88177; AAB42289.1; -;  
 SQ SEQUENCE 169 AA; 20087 MW; 6977591C CRC32;  
 Query Match 66.1%; Score 37; DB 5; Length 169;  
 Best Local Similarity 87.5%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KFKFFKFFK 8  
 |||||  
 Db 96 KFKFFKFFK 103  
 RESULT 11  
 ID 017265 PRELIMINARY; PRT; 592 AA.  
 AC 017265;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE R148.5 PROTEIN.  
 GN R148.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,

RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA LE T.T., KEMP K., SCHEET P.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF025467; AAB71039.1; -;  
 SQ SEQUENCE 592 AA; 66410 MW; DED7C7E2 CRC32;

Query Match 66.1%; Score 37; DB 5; Length 592;  
 Best Local Similarity 87.5%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFFKFF 9  
 DB 36 FKFFKFF 43

RESULT 12  
 O44959 PRELIMINARY; PRT; 548 AA.  
 ID O44959  
 AC O44959;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE M01B12.5 PROTEIN.  
 GN M01B12.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA BECKER M., GRAVES T., OZERSKY P.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF067624; AAC17564.1; -;  
 DR PFAM; PF01163; RIO1.1.  
 DR PROSITE; PS01245; RIO1.1.  
 SQ SEQUENCE 548 AA; 63781 MW; DA084B86 CRC32;

Query Match 66.1%; Score 37; DB 5; Length 548;  
 Best Local Similarity 86.7%; Pred. No. 86;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 DB 264 KFYQFFVFF 272

RESULT 13  
 O18207 PRELIMINARY; PRT; 619 AA.  
 ID O18207  
 AC O18207;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE Y48E1C.3 PROTEIN.  
 GN Y48E1C.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA MCMURRAY A.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 293394; CAB07701.1; -;  
 DR PFAM; PF00566; TBC.1.  
 SQ SEQUENCE 619 AA; 72914 MW; 5A4A7D9C CRC32;

Query Match 66.1%; Score 37; DB 5; Length 619;  
 Best Local Similarity 66.7%; Pred. No. 96;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFFKFFK 10  
 DB 426 FFEFFAFFE 434

RESULT 14  
 O50972 PRELIMINARY; PRT; 148 AA.  
 ID O50972  
 AC O50972;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN BIC08.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Plasmid cp9.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE000791; AAC66306.1; -;  
 DR TIGR; BIC08; -;  
 KW Plasmid.  
 SQ SEQUENCE 148 AA; 17585 MW; DE3D19F0 CRC32;

Query Match 64.3%; Score 36; DB 2; Length 148;  
 Best Local Similarity 75.0%; Pred. No. 39;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 FKFFKFFK 10  
 DB 132 FKFKFFK 139

```

RESULT 15
P72840
ID P72840 PRELIMINARY; PRT; 409 AA.
AC P72840;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 46.6 KD PROTEIN.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC6803;
RA TABATA S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC6803;
RX MEDLINE: 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90901; BAA16855.1; -.
KW Hypothetical protein.
SQ SEQUENCE 409 AA; 46637 MW; 12A990FF CRC32;

Query Match 64.3%; Score 36; DB 2; Length 409;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9
Db 92 KLFPEFKYF 100

```

Search completed: September 7, 1999, 20:34:41  
 Job time: 19737 sec

D**b** 76 YFKFFEIEFK 84

```

RESULT 2
SVL_SYN3
ID SVL_SYN3 STANDARD; PRT; 988 AA.
AC P73505;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE ISOLEUCYL-TRNA SYNTHETASE (EC 6.1.1.5) (ISOLEUCINE--TRNA LIGASE)
DE (ILERS)
GN ILES OR SLL1362.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROOCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K.,
RA OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A.,
RA YAMADA M., YASUDA M., TABATA S.;
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: ATP + L-ISOLEUCINE + TRNA(ILE) = AMP +
CC PYROPHOSPHATE + L-ISOLEUCYL-TRNA(ILE).
CC -!- COFACTOR: BINDS A ZINC ION (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90907; G1652625; -.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PFAM; PF00133; trna-synt_1; 1.
KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING;
KW METAL-BINDING; ZINC.
FT SIMILAR 60 70 "HIGH" REGION.
FT SIMILAR 611 615 "KMSKS" REGION.
FT BINDING 614 614 ATP (BY SIMILARITY).
SQ SEQUENCE 988 AA; 112345 MW; C89678C2 CRC32;

Query Match 66.1%; Score 37; DB 1; Length 988;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 FFFKFFK 10
Db 724 FFFKFFK 731

RESULT 3
YCX3_EUGGR
ID YCX3_EUGGR STANDARD; PRT; 634 AA.
AC P31916; P31917;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 78.0 KD PROTEIN IN PSBC INTRON 2 (ORF635).
OS EUGLENA GRACILIS.
OG CHLOROPLAST.
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE; 93347989.
RA HALLICK R.B., HONG L., DRAGER R.G., FAVREAU M.R., MONFORT A.,
RA ORSAT B., SPIELMANN A., STUTZ E.;
"Complete sequence of Euglena gracilis chloroplast DNA.";
RL NUCLEIC ACIDS RES. 21:3537-3544(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11874; -. NOT_ANNOTATED_CDS.
DR EMBL; X70810; G415736; -.
DR PIR; S34499; S34499.
DR PIR; S34500; S34500.
KW CHLOROPLAST; HYPOTHETICAL PROTEIN.
SQ SEQUENCE 634 AA; 78049 MW; 9EE16D0D CRC32;

Query Match 66.1%; Score 37; DB 1; Length 634;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 FFFKFFK 9
Db 491 FFFKFFK 497

RESULT 4
DPO1_AQUAE
ID DPO1_AQUAE STANDARD; PRT; 574 AA.
AC O67779;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE I (EC 2.7.7.7) (POL I).
GN POLA.
OS AQUIFEX ABOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
"The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL NATURE 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000765; G2984205; -.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION; DNA REPAIR;
KW HYDROLASE; EXONUCLEASE; DNA-BINDING.
SQ SEQUENCE 574 AA; 65833 MW; 95482A2F CRC32;
```



Query Match 64.3%; Score 36; DB 1; Length 574;  
Best Local Similarity 70.0%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
:|||||  
DB 444 RFFKFKRAFK 453

RESULT 5  
HS30\_YEAST STANDARD; PRT; 332 AA.  
AC P25619: Q04556;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE 30 KD HEAT SHOCK PROTEIN.  
GN HSP30 OR YCR021C OR YCR21C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
[1]  
SEQUENCE FROM N.A.  
RN STRAIN-SK02N;  
RC MEDLINE; 93306747.  
RA REGNAQ M., BOUCHERIE H.;  
RT "Isolation and sequence of HSP30, a yeast heat-shock gene coding for  
a hydrophobic membrane protein.";  
RL CURR. GENET. 23:435-442(1993).  
[2]  
SEQUENCE FROM N.A.  
RN RA FELDMANN H., MANNAUPT G., VETTER I.;  
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: PROBABLY COOPERATES WITH OTHER HEAT SHOCK PROTEINS IN  
IN THE TRANSLLOCATION OF POLYPEPTIDES THROUGH MEMBRANES. IT MAY  
CC COUNTERACT THE ALTERING EFFECT OF HEAT SHOCK ON THE PLASMA  
CC MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE ENTRY INTO STATIONARY  
CC PHASE RESULTING FROM YEAST GROWTH LIMITATION.  
CC -1- SIMILARITY: TO YEAST YRO2

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M93123; G405821; -;  
CC EMBL; X59720; E264485; -;  
CC PIR; S30781; S30781.  
CC PIR; S19432; S19432.  
CC SGD; L0000818; HSP30.  
CC HEAT SHOCK; TRANSMEMBRANE.  
FT TRANSMEM 35 55 POTENTIAL.  
FT TRANSMEM 66 86 POTENTIAL.  
FT TRANSMEM 122 142 POTENTIAL.  
FT TRANSMEM 158 178 POTENTIAL.  
FT TRANSMEM 182 202 POTENTIAL.  
FT TRANSMEM 216 236 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT DOMAIN 290 332 GLU-RICH (ACIDIC).  
SQ SEQUENCE 332 AA; 37044 MW; FCD9C61B CRC32;

Query Match 64.3%; Score 36; DB 1; Length 332;  
Best Local Similarity 75.0%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFFKFFKFF 9  
:|||||

DB 76 FFFKFFKFF 83

RESULT 6  
Y302\_MYCGE STANDARD; PRT; 317 AA.  
AC P47544; 1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN MG302.  
GN MG302  
OS MYCOPLASMA GENITALIUM.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLICUTES;  
OC MYCOPLASMATACEAE; MYCOPLASMA.  
[1]  
SEQUENCE FROM N.A.  
RN STRAIN-ATCC 33530 / G-37;  
RC MEDLINE; 96026346.  
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,  
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,  
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,  
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,  
RA TOMB J.F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,  
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium.";  
RL SCIENCE 270:397-403(1995).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U39711; G1046001; -;  
CC TIGR; MG302; -;  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 18 38 POTENTIAL.  
FT TRANSMEM 58 78 POTENTIAL.  
FT TRANSMEM 92 112 POTENTIAL.  
FT TRANSMEM 130 150 POTENTIAL.  
FT TRANSMEM 159 179 POTENTIAL.  
FT TRANSMEM 202 222 POTENTIAL.  
FT TRANSMEM 252 272 POTENTIAL.  
SQ SEQUENCE 317 AA; 36667 MW; EF5F2980 CRC32;

Query Match 64.3%; Score 36; DB 1; Length 317;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
:|||||  
DB 152 RFFKFKLKF 160

RESULT 7  
YCV4\_YEAST STANDARD; PRT; 136 AA.  
AC P25640;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 15.8 KD PROTEIN IN SUP61-RAD18 INTERGENIC REGION.  
GN YCR064C OR YCR64C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
[1]  
SEQUENCE FROM N.A.

RA ANTOINE G., BENIT P., CHANET R., FABRE R., FAYE G., FUKUHARA H.,  
 RA MATHIEU A., SOR F.;  
 RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X59720; E264407; -  
 DR PIR; S19479; S19479.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 136 AA; 15776 MW; 8FACF40B CRC32;

Query Match 64.3%; Score 36; DB 1; Length 136;  
 Best Local Similarity 66.7%; Pred. No. 6.8;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 |||||:|  
 Db 88 KFFFFQF 96

RESULT 8  
 SRE2 CAEEL STANDARD; PRT; 616 AA.  
 AC Q09273;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE SRE-2 PROTEIN.  
 GN SRE-2 OR CALCA 2.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODITINAE; CAENORHABDITIS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BURTON J.;  
 RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRE  
 CC FAMILY.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; 248045; E1344654; -  
 DR WORMPEP; C41C4.2; C601517.  
 KW TRANSMEMBRANE; MULTIGENE FAMILY.  
 FT TRANSMEM 7 27  
 FT TRANSMEM 129 149 POTENTIAL.  
 FT TRANSMEM 179 199 POTENTIAL.  
 FT TRANSMEM 269 289 POTENTIAL.  
 FT TRANSMEM 303 323 POTENTIAL.  
 FT TRANSMEM 396 416 POTENTIAL.  
 FT TRANSMEM 426 446 POTENTIAL.  
 FT TRANSMEM 482 502 POTENTIAL.  
 FT TRANSMEM 522 542 POTENTIAL.  
 SQ SEQUENCE 616 AA; 70480 MW; 2FC75724 CRC32;

Query Match 62.5%; Score 35; DB 1; Length 616;  
 Best Local Similarity 75.0%; Pred. No. 40;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FFKFFKFF 9  
 |||||:|  
 Db 4 FFKFFSIF 11

RESULT 9  
 Y453\_HAEIN STANDARD; PRT; 174 AA.  
 ID Y453\_HAEIN  
 AC P43999;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN HI0453.  
 GN HI0453.  
 OS HAEMOPHILUS INFLUENZAE.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;  
 OC HAEMOPHILUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20;  
 RX MEDLINE; 95350630.  
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
 RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOW E., COTTON M.D.,  
 RA UTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
 RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,  
 RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd".  
 RL SCIENCE 269:496-512(1995).

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; U32728; G1573438; -  
 DR TIGR; HI0453; -  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 8 28 POTENTIAL.  
 FT TRANSMEM 146 166 POTENTIAL.  
 SQ SEQUENCE 174 AA; 20416 MW; 23634093 CRC32;

Query Match 62.5%; Score 35; DB 1; Length 174;  
 Best Local Similarity 75.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 8  
 |||||:|  
 Db 3 KFFKFFLE 10

RESULT 10  
 YCHO\_YEAST STANDARD; PRT; 615 AA.  
 AC P25596; P25597; P25599;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 68.9 KD PROTEIN IN HML 5' REGION.  
 GN YCL073C OR YCL70C/71C/73C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCETES.

RP SEQUENCE FROM N.A.  
RA VAN DER AART Q.J.N., STEENMA H.Y.;  
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RN REVISIONS.  
RA GROWADKA R.;  
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE YC170C/YHL047C/YKR106W FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X59720; E264591; -;  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 121 141 POTENTIAL.  
FT TRANSMEM 153 173 POTENTIAL.  
FT TRANSMEM 187 207 POTENTIAL.  
FT TRANSMEM 217 237 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 308 328 POTENTIAL.  
FT TRANSMEM 344 364 POTENTIAL.  
FT TRANSMEM 384 404 POTENTIAL.  
FT TRANSMEM 408 428 POTENTIAL.  
FT TRANSMEM 441 461 POTENTIAL.  
FT TRANSMEM 472 492 POTENTIAL.  
FT TRANSMEM 548 568 POTENTIAL.  
SQ SEQUENCE 615 AA; 68898 MW; 0BBD5B37 CRC32;  
-----  
Query Match 62.5%; Score 35; DB 1; Length 615;  
Best Local Similarity 75.0%; Pred. NO. 40;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 EFKFKFF 9  
DB 353 FNFFTFF 360  
-----  
RESULT 11  
ID YK86\_YEAST STANDARD; PRT; 615 AA.  
AC P36173;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 69.0 KD PROTEIN IN SIR1 3'REGION.  
GN YKR106W.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GAILLON L., DUJON B.;  
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE YC170C/YHL047C/YKR106W FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: Z28202; G486617; -;  
PIR: S40624; S40624.  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 121 141 POTENTIAL.  
FT TRANSMEM 153 173 POTENTIAL.  
FT TRANSMEM 217 237 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 308 328 POTENTIAL.  
FT TRANSMEM 344 364 POTENTIAL.  
FT TRANSMEM 384 404 POTENTIAL.  
FT TRANSMEM 408 428 POTENTIAL.  
FT TRANSMEM 441 461 POTENTIAL.  
FT TRANSMEM 472 492 POTENTIAL.  
FT TRANSMEM 549 569 POTENTIAL.  
SQ SEQUENCE 615 AA; 68954 MW; A8423CFB CRC32;  
-----  
Query Match 62.5%; Score 35; DB 1; Length 615;  
Best Local Similarity 75.0%; Pred. NO. 40;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 EFKFKFF 9  
DB 353 FNFFTFF 360  
-----  
RESULT 12  
ID Y093\_CAEEL STANDARD; PRT; 504 AA.  
AC P41843;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 58.5 KD PROTEIN T20B12.3 IN CHROMOSOME III.  
GN T20B12.3.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U10401; G500716; -;  
DR WORMPEP; T20B12.3; CE01409.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 504 AA; 58514 MW; BCCDDBB9 CRC32;  
-----  
Query Match 62.5%; Score 35; DB 1; Length 504;  
Best Local Similarity 60.0%; Pred. NO. 34;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KFKFKFFK 10  
DB 66 KYFKTIKFF 75  
-----  
RESULT 13  
ID Y0T5\_CAEEL STANDARD; PRT; 824 AA.  
AC P34651;  
DT 01-FEB-1994 (REL. 28, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 95.7 KD PROTEIN ZK632.5 IN CHROMOSOME III.  
 GN ZK632.5.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDAE;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,  
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,  
 RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,  
 RA LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,  
 RA SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,  
 RA WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,  
 RA WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans".  
 RL NATURE 368:32-38(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z22181; E1351189; -  
 DR PIR; S40937; S40937.  
 DR WORMPEP; ZK632.5; CE00422.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 824 AA; 95726 MW; F3506E80 CRC32;  
 Query Match 62.5%; Score 35; DB 1; Length 824;  
 Best Local Similarity 66.7%; Pred. No. 53;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKFKFKFF 9  
 |||||  
 Db 553 KFEFLKLF 561  
 RESULT 14  
 GLBC\_NIPBR STANDARD; PRT; 173 AA.  
 AC P51536;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE GLOBIN, CUTICULAR ISOFORM PRECURSOR.  
 GN GLBC.  
 OS NIPPOSTRONGYLUS BRASILIENSIS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; STRONGYLIDA;  
 OC TRICHOSTRONGYLOIDEA; HELIGMONELLIDAE; NIPPOSTRONGYLINAE;  
 OC NIPPOSTRONGYLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE; 95198760.  
 RA BLAXTER M.L., INGRAM L., TWEEDIE S.;  
 RT "Sequence, expression and evolution of the globins of the parasitic  
 RT nematode Nippostrongylus brasiliensis".  
 RL MOL. BIOCHEM. PARASITOL. 68:1-14(1994).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY BY ADULT NEMATODES IN THE GUT.  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR (BY SIMILARITY).  
 CC -!- THE GLOBINS OF THE NEMATODE PARASITE N. BRASILIENSIS HAVE OXYGEN

CC AFFINITIES 100-FOLD HIGHER THAN THE RODENT HOST'S HEMOGLOBINS.  
 CC TWO ISOFORMS ARE FOUND, ONE LOCATED IN THE CUTICLE, AND THE OTHER  
 CC IN THE BODY OF THE NEMATODE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L20896; G309617; -  
 DR EMBL; L25872; G414596; -  
 DR PROSITE; PS01033; GLOBIN; 1.  
 DR PFAM; PF00042; Globin; 1.  
 DR HSP; P28316; IASH.  
 KW HEME; OXYGEN TRANSPORT; RESPIRATORY PROTEIN; SIGNAL.  
 FT SIGNAL 1 16  
 FT CHAIN 17 173  
 FT METAL 113 113 GLOBIN, CUTICULAR ISOFORM.  
 FT TRON (HEME PROXIMAL LIGAND)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 173 AA; 19696 MW; 9A473C5D CRC32;  
 Query Match 60.7%; Score 34; DB 1; Length 173;  
 Best Local Similarity 47.1%; Pred. No. 18;  
 Matches 8; Conservative 1; Mismatches 0; Indels 8; Gaps 1;  
 QY 2 FPKFF-----KPK 10  
 |||||  
 Db 46 FYKFFTHKDLRKKF 62  
 RESULT 15  
 ING\_CERTO STANDARD; PRT; 165 AA.  
 ID ING\_CERTO  
 AC P42162;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE INTERFERON GAMMA PRECURSOR (IFN-GAMMA).  
 GN IFNG.  
 OS CERCOCEBUS TORQUATUS ATYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; CERCOPTHECIDAE; CERCOPTHECINAE; CERCOCEBUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96003435.  
 RA VILLINGER F.J., BRAR S.S., MAYNE A.E., CHIKKALA N., ANSARI A.A.;  
 RT "Comparative sequence analysis of cytokine genes from human and  
 RT nonhuman primates".  
 RL J. IMMUNOL. 155:3946-3954(1995).  
 CC -!- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS  
 CC OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,  
 CC HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR  
 CC OF MACROPHAGES, IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED  
 CC CELLS, & IT CAN POTENTIATE THE ANTIVIRAL AND ANTI-TUMOR EFFECTS OF  
 CC THE TYPE I INTERFERONS.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- TISSUE SPECIFICITY: RELEASED PRIMARILY FROM ACTIVATED T  
 CC LYMPHOCYTES.  
 CC -!- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L26025; G514336; -

DR PFAM: PF00714; IFN-gamma; 1.  
DR HSP: P07353; IRFB.  
KW CYTOKINE: ANTIVIRAL; GROWTH REGULATION; IMMUNOMODULATION;  
KW GLYCOPROTEIN; SIGNAL.  
FT SIGNAL 1 20 BY SIMILARITY.  
FT CHAIN 21 165 INTERFERON GAMMA.  
FT CARBOHYD 48 48 POTENTIAL.  
FT CARBOHYD 120 120 POTENTIAL.  
SQ SEQUENCE 165 AA; 19301 MW; F46865C7 CRC32;

Query Match 60.7%; Score 34; DB 1; Length 165;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFFK 10  
: || || ||  
Db 76 YFKLKRSEK 84

Search completed: September 7, 1999, 23:50:13  
Job time: 1952 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:24 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-10  
Perfect score: 56  
Sequence: 1 KFFKFFKFFK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	39	69.6	2010	2	B71616	phosphatase (acid
2	38	67.9	873	2	S53828	cytochrome-c oxida
3	38	67.9	166	2	S10513	interferon gamma -
4	38	67.9	1307	2	S53412	probable membrane
5	37	66.1	988	2	S77211	isoleucine--trNA 1
6	37	66.1	192	2	H70138	cell division prot
7	37	66.1	241	2	S34500	hypothetical prote
8	36	64.3	148	2	A70221	conserved hypothet
9	36	64.3	409	2	S74704	hypothetical prote
10	36	64.3	317	2	D64233	hypothetical prote
11	36	64.3	411	2	T01471	hypothetical prote
12	36	64.3	332	2	S31848	heat shock protein
13	36	64.3	136	2	S19479	hypothetical prote
14	36	64.3	2206	2	G71611	hypothetical prote
15	36	64.3	1014	2	H71602	protein with DnaJ
16	36	64.3	574	2	G70468	DNA polymerase I (
17	35	62.5	174	2	H64007	hypothetical prote
18	35	62.5	615	2	S74277	probable membrane
19	35	62.5	615	2	S40624	probable membrane
20	35	62.5	240	2	A64652	hypothetical prote
21	35	62.5	240	2	D71939	hypothetical prote
22	35	62.5	824	2	S40937	hypothetical prote
23	34	60.7	1082	1	RNEGB	DNA-directed RNA p
24	34	60.7	1386	1	RNLVC2	DNA-directed RNA p
25	34	60.7	193	2	S07734	NADH dehydrogenase
26	34	60.7	516	2	S34525	hypothetical prote
27	34	60.7	649	2	G71530	probable ATP synth
28	34	60.7	117	2	S66720	probable membrane
29	34	60.7	2339	2	A45597	DNA-directed RNA p
30	33	58.9	166	1	IVHUG	interferon gamma p
31	33	58.9	166	1	IVCJG	interferon gamma p
32	33	58.9	902	2	S34495	probable carrier p
33	33	58.9	144	2	B30471	hypothetical prote
34	33	58.9	569	2	A45624	trophozoite cystei
35	33	58.9	459	2	F70189	rod shape-determin
36	33	58.9	233	2	E54596	ABC transporter, A
37	33	58.9	376	2	G71828	hypothetical prote
38	33	58.9	382	2	E54686	myosin-like protei
39	33	58.9	233	2	E71950	probable ABC trans

ALIGNMENTS

RESULT 1

B71616  
phosphatase (acid phosphatase family) PFB0380c - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: B71616  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H. Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: B71616  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2010 <GAR>  
A:Cross-references: GB:AE001391; GB:AE001362; NID:g3845168; PID:g3845169; TIGR:PFB038  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0380c

Query Match 69.6%; Score 39; DB 2; Length 2010;  
Best Local Similarity 87.5%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFFKFFKFFK 8  
|||||  
Db 801 KFFKLFKFF 808

RESULT 2

S53828  
cytochrome-c oxidase (EC 1.9.3.1) chain 1 and chain 2 - Acanthamoeba castellanii mito  
C:Species: mitochondrion Acanthamoeba castellanii  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 29-Jan-1999  
C:Accession: S53828; S68552  
R:Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W. J. Mol. Biol. 245, 522-537, 1995  
A:Title: The mitochondrial DNA of the amoeboid protozoon, Acanthamoeba castellanii: c  
A:Reference number: S53825; MUID:9514725  
A:Accession: S53828  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-873 <BUR>  
A:Cross-references: GB:U12386; NID:g562028; PID:g562032  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
R:Loneragan, K.M.; Gray, M.W. J. Mol. Biol. 257, 1019-1030, 1996  
A:Title: Expression of a continuous open reading frame encoding subunits 1 and 2 of c  
A:Reference number: S68552; MUID:96192086  
A:Accession: S68552  
A:Molecule type: DNA  
A:Residues: 1-873 <LON>  
A:Cross-references: GB:U12386; NID:g562028; PID:g562032  
A:Experimental source: ATCC 30010  
C:Genetics:  
A:Gene: cox1/2  
A:Genome: mitochondrion  
A:Genetic code: SCC6  
C:Superfamily: Acanthamoeba cytochrome-c oxidase chain I/II; cytochrome-c oxidase cha  
C:Keywords: copper binding; electron transfer; heme; membrane-associated complex; mit  
F:23-470/Domain: cytochrome-c oxidase chain I homology <CO1>

F:568-854/Domain: cytochrome-c oxidase chain II homology <CO2>

Query Match 67.9%; Score 38; DB 2; Length 873;  
Best Local Similarity 87.5%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKEFFK 9  
|||||  
Db 530 FKFKEFFK 537

RESULT 3

S10513  
interferon gamma - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Sep-1997  
C:Accession: S10513  
R:Diikmans, R.; Vandenbroeck, K.; Beuken, E.; Billiau, A.  
Nucleic Acids Res. 18, 4259, 1990  
A:Title: Sequence of the porcine interferon-gamma (IFN-gamma) gene.  
A:Reference number: S10513; MUID:90332433  
A:Accession: S10513  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-166 <DIJ>  
A:Cross-references: EMBL:X53085; NID:g1979; PID:g927562  
C:Genetics:  
A:Introns: 38/3; 61/3; 123/3  
C:Superfamily: interferon gamma

Query Match 67.9%; Score 38; DB 2; Length 166;  
Best Local Similarity 66.7%; Pred. No. 5; 6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKEFFK 10  
:||||:  
Db 76 YKFKEFFK 84

RESULT 4

S53412  
probable membrane protein YLR425w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein L9576.5  
C:Species: Saccharomyces cerevisiae  
C:Date: 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 04-Sep-1998  
C:Accession: S53412  
R:Favella, A.  
Submitted to the EMBL Data Library, February 1995  
A:Description: The sequence of S. cerevisiae cosmid 9576.  
A:Reference number: S53409  
A:Accession: S53412  
A:Molecule type: DNA  
A:Residues: 1-1307 <FAV>  
A:Cross-references: EMBL:U20939; NID:g664871; PID:g664876; MIPS:YLR425w  
A:Experimental source: strain S288C (AB972)  
C:Genetics:  
A:Map position: 12R  
C:Superfamily: yeast probable membrane protein YLR425w  
C:Keywords: transmembrane protein  
F:382-398/Domain: transmembrane #status predicted <TM>

Query Match 67.9%; Score 38; DB 2; Length 1307;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKFKEFFK 10  
:||||:  
Db 1075 KYKFKEFFK 1084

RESULT 5

S77211  
isoleucine--tRNA ligase (EC 6.1.1.5) - Synecocystis sp. (strain PCC 6803)  
N:Alternate names: isoleucyl-tRNA synthetase; protein sl11362  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S77211  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis S.

A:Reference number: S74322; MUID:97061201  
A:Accession: S77211  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-988 <KAN>  
A:Cross-references: EMBL:D90907; GB:AB001339; NID:g1652618; PID:d1018278; PID:g165262  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: iles  
A:Start codon: GTG  
C:Superfamily: isoleucine--tRNA ligase  
C:Keywords: ligase

Query Match 66.1%; Score 37; DB 2; Length 988;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KKFKEFFK 10  
:||||:  
Db 724 KKFKEFFK 731

RESULT 6

H70138  
cell division protein (ftsJ) homolog - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 04-Sep-1998  
C:Accession: H70138  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: H70138  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-192 <KLE>  
A:Cross-references: GB:AE001138; GB:AE000783; NID:g2688210; PID:g2688216; TIGR:BB0313  
A:Experimental source: strain B31  
C:Superfamily: cell division protein J

Query Match 66.1%; Score 37; DB 2; Length 192;  
Best Local Similarity 70.0%; Pred. No. 9.3;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKFKEFFK 10  
|||:||||  
Db 160 KKFKEFFK 169

RESULT 7

S34500  
hypothetical protein 241 (psbc 3' region) - Euglena gracilis chloroplast  
C:Species: chloroplast Euglena gracilis  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Apr-1995  
C:Accession: S34500; S34868



R;Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielmann, submitted to the EMBL Data Library, January 1993  
 A:Description: The complete sequence of the Euglena gracilis chloroplast genome (tentative)  
 A:Reference number: S34494  
 A:Accession: S34500  
 A:Molecule type: DNA  
 A:Residues: 1-241 <HAL1>  
 A:Cross-references: EMBL:X70810  
 R;Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielmann Nucleic Acids Res. 21, 3537-3544, 1993  
 A:Title: Complete sequence of Euglena gracilis chloroplast DNA.  
 A:Reference number: S34862; MUID:93347989  
 A:Accession: S34868  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-241 <HAL2>  
 A:Cross-references: EMBL:X70810  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 66.1%; Score 37; DB 2; Length 241;  
 Best Local Similarity 85.7%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 FKFEKFF 9  
 |||||  
 Db 98 FRFEKFF 104

## RESULT 8

A70221  
 conserved hypothetical protein BBC08 - Lyme disease spirochete plasmid C/cp9  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: A70221  
 R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943

A:Accession: A70221  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-148 <KLE>  
 A:Cross-references: GB:AE000791; MUID:g2689914; PID:g2689920; TIGR:BBC08  
 A:Experimental source: strain B31  
 C:Genetics:  
 A:Genome: plasmid

Query Match 64.3%; Score 36; DB 2; Length 148;  
 Best Local Similarity 75.0%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 FKFEKFF 10  
 |||||  
 Db 132 FKFEKFF 139

## RESULT 9

S74704  
 hypothetical protein slr1302 - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
 C:Accession: S74704  
 R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S74704  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-409 <KAN>  
 A:Cross-references: EMBL:D90901; GB:AB001339; MUID:g1651897; PID:d1017588; PID:g165192  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Start codon: GTG

Query Match 64.3%; Score 36; DB 2; Length 409;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 9  
 |||||  
 Db 92 KLFPFKYF 100

## RESULT 10

D64233  
 hypothetical protein MG302 - Mycoplasma genitalium (SGC3)  
 C:Species: Mycoplasma genitalium  
 C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 10-Oct-1997  
 C:Accession: D64233  
 R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R. M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C. Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346  
 A:Accession: D64233  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-317 <TIGR>  
 A:Cross-references: GB:U39711; GB:I43967; MUID:g1045997; PID:g1046001; TIGR:MG302  
 A:Experimental source: strain G-37  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 64.3%; Score 36; DB 2; Length 317;  
 Best Local Similarity 66.7%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 9  
 |||||  
 Db 152 RFFKFLKLF 160

## RESULT 11

T01471  
 hypothetical protein T24H24.4 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Mar-1999  
 C:Accession: T01471  
 R;Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K. submitted to the EMBL Data Library, August 1998  
 A:Description: The sequence of A. thaliana T24H24.  
 A:Reference number: 214333  
 A:Accession: T01471  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-411 <COU>  
 A:Cross-references: EMBL:AF075598; MUID:g3293581; PID:g3377829  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 288/1

A.Note: T24H24.4

Query Match 64.3%; Score 36; DB 2; Length 411;  
Best Local Similarity 75.0%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
||:|:| ||  
Db 127 FFEFTTF 134

#### RESULT 12

S31848  
heat shock protein HSP30 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YCR021c  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 06-Feb-1998  
C:Accession: S31848; S19432; S30781  
R:Regnacq, M.; Boucherie, H.  
Curr. Genet. 23, 435-442, 1993  
A:Title: Isolation and sequence of HSP30, a yeast heat-shock gene coding for a hydrophobic protein  
A:Reference number: S31848; MUID:93306747  
A:Accession: S31848  
A:Molecule type: DNA  
A:Residues: 1-263 <REG1>  
A:Cross-references: EMBL:M93123  
R:Feldmann, H.; Mannhaupt, G.; Vetter, I.  
submitted to the Protein Sequence Database, March 1992  
A:Reference number: S19429  
A:Accession: S19432  
A:Molecule type: DNA  
A:Residues: 1-190, 'A', 192-332 <PEL>  
A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264485; PID:g1907165; MIPS:YCR021c  
R:Regnacq, M.; Boucherie, H.  
submitted to the EMBL Data Library, January 1993  
A:Reference number: S30781  
A:Accession: S30781  
A:Molecule type: DNA  
A:Residues: 1-152, 'GY', 164-165, 167, 'A', 169, 'NSNRGVL', 170-240, 'VFNOT', 270, 278-280, 'FMVFWI'  
A:Cross-references: EMBL:M93123  
A>Note: the difference at the carboxyl end is due to a frameshift error  
C:Genetics:  
A:Gene: SGD:HSP30  
A:Cross-references: SGD:S0000615; MIPS:YCR021c  
A:Map position: 3R  
C:Keywords: membrane protein

Query Match 64.3%; Score 36; DB 2; Length 332;  
Best Local Similarity 75.0%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
||:|:| ||  
Db 76 FFEFFAFF 83

#### RESULT 13

S19479  
hypothetical protein YCR064c - yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 14-Nov-1997  
C:Accession: S19479  
R:Antoine, G.; Benit, P.; Chanet, R.; Fabre, R.; Faye, G.; Fukuhara, H.; Mathieu, A.; Schmitt, J.  
submitted to the Protein Sequence Database, March 1992  
A:Reference number: S19477  
A:Accession: S19479  
A:Molecule type: DNA  
A:Residues: 1-136 <ANT>  
A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264407; PID:g1907204; MIPS:YCR064c  
C:Genetics:  
A:Map position: 3R

Query Match 64.3%; Score 36; DB 2; Length 136;  
Best Local Similarity 66.7%; Pred. No. 9.8;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
||:|:| ||:|  
Db 88 KFFFFQYF 96

#### RESULT 14

G71611  
hypothetical protein PFB0560w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: G71611  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: G71611  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2206 <GAR>  
A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PID:g3845217; TIGR:PFB0560w  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0560w

Query Match 64.3%; Score 36; DB 2; Length 2206;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFFK 10  
||:|:| ||  
Db 1015 FKIFYAFFK 1023

#### RESULT 15

H71602  
protein with DnaJ domain (RESA-like) PFB0920w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 16-Dec-1998  
C:Accession: H71602  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: H71602  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1014 <GAR>  
A:Cross-references: GB:AE001425; GB:AE001362; NID:g3845310; PID:g3845311; TIGR:PFB0920w  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0920w  
C:Superfamily: dnaJ amino-terminal homology  
F:677-742/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 64.3%; Score 36; DB 2; Length 1014;  
Best Local Similarity 66.7%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
||:|:| ||:|  
Db 2 KYFKKFKYF 10

Search completed: September 7, 1999, 23:06:25  
Job time: 2473 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:34 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-10  
Perfect score: 56  
Sequence: 1 KFFKFFKFFK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	10	1	US-08-097-830E-10
2	56	100.0	10	2	US-08-456-112B-10
3	51	91.1	9	2	US-08-456-112B-40
4	38	67.9	11	1	US-08-049-871-8
5	38	67.9	11	1	US-07-819-893-8
6	38	67.9	11	1	US-08-280-397-8
7	38	67.9	10	1	US-08-097-830E-8
8	38	67.9	12	1	US-08-097-830E-20
9	38	67.9	10	2	US-08-456-112B-8
10	38	67.9	12	2	US-08-456-112B-20
11	38	67.9	11	2	US-08-456-112B-37
12	33	58.9	146	1	US-08-030-077-1
13	33	58.9	122	1	US-08-354-456A-9
14	33	58.9	143	2	US-08-449-639-8
15	33	58.9	10	2	US-08-456-112B-41
16	32	57.1	3135	1	US-08-323-170B-2
17	32	57.1	880	2	US-08-916-917-12
18	31	55.4	887	1	US-07-867-106-3
19	31	55.4	463	1	US-08-426-428-2
20	31	55.4	350	1	US-08-415-751-16
21	31	55.4	1089	1	US-08-375-300-2
22	31	55.4	764	1	US-08-375-300-4
23	31	55.4	1089	3	PCT-US95-16930-2
24	31	55.4	764	3	PCT-US95-16930-4
25	30	53.6	1264	1	US-07-789-915A-6
26	30	53.6	775	1	US-07-603-133B-14
27	30	53.6	8	1	US-08-257-782-25
28	30	53.6	1264	1	US-08-005-002C-6
29	30	53.6	920	1	US-08-101-593-2
30	30	53.6	920	1	US-08-101-593-4
31	30	53.6	1264	1	US-08-487-203A-6
32	30	53.6	8	2	US-08-577-846-25
33	30	53.6	78	2	US-08-680-726A-60
34	30	53.6	310	2	US-08-695-355-4
35	30	53.6	171	2	US-08-665-040-2
36	29	51.8	525	1	US-08-077-939-19
37	29	51.8	535	1	US-08-461-599-19
38	29	51.8	533	1	US-08-488-305A-6
39	29	51.8	525	1	US-08-461-621-19

40 29 51.8 525 1 US-08-465-334-19 Sequence 19, Appl  
41 29 51.8 208 2 US-08-462-169B-20 Sequence 20, Appl  
42 29 51.8 1434 2 US-08-540-406-10 Sequence 10, Appl  
43 29 51.8 1447 2 US-08-540-406-19 Sequence 19, Appl  
44 29 51.8 877 2 US-08-916-917-2 Sequence 2, Appl  
45 29 51.8 877 2 US-08-972-631-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-10  
; Sequence 10, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-10  
  
Query Match 100.0%; Score 56; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KFFKFFKFFK 10  
| | | | | | | | | |  
Db 1 KFFKFFKFFK 10  
  
RESULT 2  
US-08-456-112B-10  
; Sequence 10, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-10

Query Match 100.0%; Score 56; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00057;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
Db 1 KFFKFFKFFK 10

RESULT 3  
US-08-456-112B-40  
Sequence 40, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-40

Query Match 91.1%; Score 51; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e-04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
Db 1 KFFKFFKFF 9

RESULT 4  
US-08-049-871-8  
Sequence 8, Application US/08049871  
Patent No. 5358933  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/049,871  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/658,744  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-049-871-8

Query Match 67.9%; Score 38; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
Db 2 KFLKFLKFLK 11

RESULT 5  
US-07-819-893-8  
; Sequence 8, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-8

Query Match 67.9%; Score 38; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KFFKFFKFK 10  
Db 2 KFLKFLKFLK 11

RESULT 6  
US-08-280-397-8  
; Sequence 8, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-280-397-8

Query Match 67.9%; Score 38; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KFFKFFKFK 10  
Db 2 KFLKFLKFLK 11

RESULT 7  
US-08-097-830E-8  
; Sequence 8, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
US-08-097-830E-8

Query Match          67.9%; Score 38; DB 1; Length 10;
Best Local Similarity 70.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10
DB 1 KFLKFLKFLK 10

RESULT 8
US-08-097-830E-20
; Sequence 20, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097.830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; OTHER INFORMATION: sulfide bond between Cys and Cys
US-08-097-830E-20

Query Match          67.9%; Score 38; DB 1; Length 12;
Best Local Similarity 70.0%; Pred. No. 0.44;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10
DB 2 KFLKFLKFLK 11

RESULT 9
US-08-456-112B-8
; Sequence 8, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
```

```
;
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-8

Query Match          67.9%; Score 38; DB 2; Length 10;
Best Local Similarity 70.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10
DB 1 KFLKFLKFLK 10

RESULT 10
US-08-456-112B-20
; Sequence 20, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```



ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-20

Query Match 67.9%; Score 38; DB 2; Length 12;  
Best Local Similarity 70.0%; Pred. No. 0.44;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
DB 2 KFLKFLKFLK 11

RESULT 11  
US-08-456-112B-37  
Sequence 37, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-37

Query Match 67.9%; Score 38; DB 2; Length 11;  
Best Local Similarity 70.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKFKFKFK 10  
DB 2 KFLKFLKFLK 11

RESULT 12  
US-08-030-077-1  
Sequence 1, Application US/08030077  
Patent No. 5451658  
GENERAL INFORMATION:  
APPLICANT: Seelig, Gail F.  
TITLE OF INVENTION: Antagonists of Human Gamma Interferon  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: One Giralda Farms  
CITY: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.5  
SOFTWARE: Microsoft Word 4.00B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/030,077  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/030,077  
FILING DATE: 19-March-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: JB0151K  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
TELEX: 219165  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 146 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-030-077-1

Query Match 58.9%; Score 33; DB 1; Length 146;  
Best Local Similarity 66.7%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFKFK 10  
DB 56 YFKLFKNFK 64

RESULT 13  
US-08-354-456A-9  
Sequence 9, Application US/08354456A  
Patent No. 5567611  
GENERAL INFORMATION:  
APPLICANT: Ralph, Peter  
APPLICANT: Martin, George  
APPLICANT: Platek, Michael  
APPLICANT: Larrick, James W.  
TITLE OF INVENTION: Multifunctional M-CSF Proteins and Genes Encoding  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHIRON CORPORATION  
STREET: Intellectual Property - R440, P.O. Box 8097  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.

ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/354,456A  
FILING DATE: 12-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/995,338  
FILING DATE: 21-DEC-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McGarrigle Jr., Philip L.  
REGISTRATION NUMBER: 31,395  
REFERENCE/DOCKET NUMBER: 750.003/32387  
TELEPHONE: (510) 601-2718  
TELEFAX: (510) 655-3542  
TELEX: n/a  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-354-456A-9

Query Match 58.9%; Score 33; DB 1; Length 122;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFFK 10  
: || || ||  
Db 51 YFKLFKNFK 59

RESULT 14  
US-08-449-639-8  
Sequence 8, Application US/08449639  
Patent No. 5770191  
GENERAL INFORMATION:  
APPLICANT: Johnson, Howard M.  
APPLICANT: Szente, Brian E.  
TITLE OF INVENTION: C-Terminal Peptides of Interferon-Gamma  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,639  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Saliwanchik, David R.  
REGISTRATION NUMBER: 31,794  
REFERENCE/DOCKET NUMBER: UF150  
TELEPHONE: 904-375-8100  
TELEFAX: 904-375-8100

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 143 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-449-639-8

Query Match 58.9%; Score 33; DB 2; Length 143;  
Best Local Similarity 66.7%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFFK 10  
: || || ||  
Db 53 YFKLFKNFK 61

RESULT 15  
US-08-456-112B-41  
Sequence 41, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-41

Query Match 58.9%; Score 33; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2.3;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFFKFFK 8  
: || || ||  
Db 2 KFKLFKFFK 9

Search completed: September 7, 1999, 22:38:35  
Job time: 7919 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:12 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-10  
Perfect score: 56  
Sequence: 1 KFFKFFKFFK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues.

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	100.0	10	1 R71781	Peptide neutralisi
2	56	100.0	10	1 W21598	Antibiotic potenti
3	45	80.4	9	1 W21628	Antibiotic potenti
4	39	69.6	9	1 P50399	IFN-gamma-(48-148)
5	39	69.6	15	1 R57402	Peptide for treati
6	38	67.9	32	1 R33514	Generic peptide fo
7	38	67.9	11	1 R33532	Peptide for treati
8	38	67.9	11	1 R39294	Endotoxin lipid A
9	38	67.9	10	1 R71779	Peptide neutralisi
10	38	67.9	12	1 R71791	Peptide neutralisi
11	38	67.9	11	1 W21625	Antibiotic potenti
12	38	67.9	10	1 W21596	Antibiotic potenti
13	38	67.9	12	1 W21608	Antibiotic potenti
14	38	67.9	166	1 W83904	Porcine interferon
15	38	67.9	166	1 W73233	Porcine interferon
16	36	64.3	18	1 P90327	Alpha-helix-formin
17	36	64.3	574	1 Y00940	Aquifex VF-5 DNA p
18	35	62.5	228	1 W98448	H. pylori GHPO 576
19	35	62.5	240	1 Y11023	H. pylori ORF 029e
20	34	60.7	17	1 R57403	Peptide for treati
21	34	60.7	15	1 R57404	Peptide for treati
22	34	60.7	12	1 R57387	Peptide for treati
23	34	60.7	323	1 W15759	Cotton fibrous tis
24	34	60.7	8	1 W11166	Peptide used in an
25	33	58.9	144	1 P91033	Modified gamma-int
26	33	58.9	147	1 P91036	IFNX 957 gene. Mod
27	33	58.9	153	1 P91039	Synthetic interfe
28	33	58.9	144	1 P91040	IFNX 918 gamma-int
29	33	58.9	147	1 P91042	Modified IFNX 930
30	33	58.9	132	1 P90362	Recombinant Group
31	33	58.9	166	1 P99990	Modified gamma-int
32	33	58.9	144	1 P90822	Analogue(s) of int
33	33	58.9	138	1 R03067	Human synthetic in
34	33	58.9	145	1 P80406	Fusion protein enc
35	33	58.9	144	1 P80404	IFN-gamma fusion p
36	33	58.9	144	1 P80405	IFN-gamma fusion p
37	33	58.9	145	1 P80407	Fusion protein enc
38	33	58.9	143	1 P81103	Sequence of interf
39	33	58.9	123	1 P80535	Fragment of gamma-
40	33	58.9	144	1 P81848	Sequence of homoge
41	33	58.9	144	1 P91024	Sequence of mature
42	33	58.9	134	1 P82569	Human Interferon-g
43	33	58.9	74	1 P81206	Left half of gamma

ALIGNMENTS

RESULT 1	
R71781	ID R71781 standard; peptide; 10 AA.
AC R71781;	AC R71781;
DT 01-OCT-1995 (first entry)	DT 01-OCT-1995 (first entry)
DE Peptide neutralising toxicity of Lipid A.	DE Peptide neutralising toxicity of Lipid A.
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.	KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.
OS Synthetic.	OS Synthetic.
PN W09503327-A.	PN W09503327-A.
PD 02-FEB-1995.	PD 02-FEB-1995.
PF 21-JUL-1994; E02413.	PF 21-JUL-1994; E02413.
PR 26-JUL-1993; US-097830.	PR 26-JUL-1993; US-097830.
PA (BIOS-) BIOSYNTH SRL.	PA (BIOS-) BIOSYNTH SRL.
PI Porro M;	PI Porro M;
DR WPI; 95-075190/10.	DR WPI; 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids	PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids
PS Claim 12; Page 21; 26pp; English.	PS Claim 12; Page 21; 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.	CC New peptides are claimed which are linear or cyclic peptides of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable CC nutrient solutions, etc.	CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new CC peptides.	CC The present sequence is a specifically claimed example of the new CC peptides.
SQ Sequence 10 AA;	SQ Sequence 10 AA;
Query Match 100.0%; Score 56; DB 1; Length 10;	Query Match 100.0%; Score 56; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00045;	Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KFFKFFKFFK 10	QY 1 KFFKFFKFFK 10
Db 1 KFFKFFKFFK 10	Db 1 KFFKFFKFFK 10
RESULT 2	
W21598	ID W21598 standard; peptide; 10 AA.
AC W21598;	AC W21598;
DT 26-AUG-1997 (first entry)	DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #10.	DE Antibiotic potentiating peptide #10.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;	KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.	KW permeability; outer bacterial membrane.
OS Synthetic.	OS Synthetic.
PN W09638163-A1.	PN W09638163-A1.
PD 05-DEC-1996.	PD 05-DEC-1996.
PF 29-MAY-1996; E02313.	PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.	PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.	PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;	PI Porro M, Varra M;
DR WPI; 97-034095/03.	DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required	PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required
PS Claim 14; Page 25; 37pp; English.	PS Claim 14; Page 25; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used CC to treat or prevent microbial infections. These peptides bind to	CC The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used CC to treat or prevent microbial infections. These peptides bind to

Left and right hal  
Clair-BamHI fragmen

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.00045;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFFKFFKFFK 10  
 |||||  
 Db 1 KFFKFFKFFK 10

## RESULT 3

W21628 ID W21628 standard; peptide; 9 AA.

AC W21628;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #40.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 44; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 80.4%; Score 45; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFFKFFKFFK 8  
 |||||  
 Db 1 KFFKFFKFFK 8

## RESULT 4

P50399 ID P50399 standard; Protein; 99 AA.

AC P50399;  
 DT 16-MAR-1992 (first entry)  
 DE IFN-gamma-(48-148).  
 KW Interferon; homo sapiens; human.  
 OS Synthetic.  
 PN J60041697-A.  
 PD 05-MAR-1985.  
 PF 15-AUG-1983; 148026.  
 PR 15-AUG-1983; JP-148026.  
 PA (ASAH ) ASAH CHEMICAL IND KK.  
 DR WPI: 85-09095/15.  
 DR N-PSDB: N50456.  
 PT Synthesis of active protein derivs. - by reacting two peptide(s)

PT one of which is synthetically formed by DNA recombination.  
 PS Disclosure; Page 17; 23pp; Japanese.

CC The sequence was prepd. from 25 synthetic oligonucleotides which  
 CC were annealed to a further 25 making up the complementary strand,  
 CC to form a ds duplex. The sequence comprises residues 48-146 of  
 CC IFN gamma. An ile residue (no. 47) was later added to the N-  
 CC terminal of the expressed peptide. A second duplex encoding  
 CC residues 1-46 of the protein was also synthesised (see N50455).  
 CC The two peptides were expressed, purified and then reacted in the  
 CC presence of a condensing agent to prepare the complete peptide.  
 CC Alternatively one of the peptides is converted to its reactive  
 CC deriv. and then reacted with the other peptide.  
 CC See also P50398 and P50400.  
 SQ Sequence 99 AA;

Query Match 69.6%; Score 39; DB 1; Length 99;  
 Best Local Similarity 77.8%; Pred. No. 2.1;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFKFFKFFK 10  
 :|||||  
 Db 9 YFKFFKFFK 17

## RESULT 5

R57402 ID R57402 standard; Protein; 15 AA.

AC R57402;  
 DT 21-MAR-1995 (first entry)  
 DE Peptide for treating diseases related to anti-DNA antibodies.  
 KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
 OS Synthetic.  
 PN J06192290-A.  
 PD 12-JUL-1994.  
 PF 18-JAN-1993; 006098.  
 PR 30-SEP-1992; JP-261821.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 94-260510/32.  
 PT A peptide and an adsorbing agent prepd. by immobilising it on a  
 PT carrier - useful for treatment of diseases related to anti-DNA  
 PT antibodies and immune complexes  
 PS Disclosure; Page 12; 14pp; Japanese.  
 CC The sequences given in R57385-413 are peptides which are all covered  
 CC by the claimed generic formula:  
 CC H-X-(A-B)n-Y-Z  
 CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
 CC B = Trp, Phe, Asn or Glu;  
 CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
 CC consisting of 2-10 residues, provided that at least one of  
 CC X or Y are present;  
 CC Z = OH or NH2; and  
 CC n = 2-5.

CC These peptides may be immobilised on a carrier in the preparation of an  
 CC absorbing agent which may be used in the treatment of diseases related  
 CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 15 AA;

Query Match 69.6%; Score 39; DB 1; Length 15;  
 Best Local Similarity 77.8%; Pred. No. 0.39;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKFFK 9  
 |||||  
 Db 2 KFFDFDFDF 10

## RESULT 6

R33514 ID R33514 standard; peptide; 32 AA.

AC R33514;  
 DT 07-JUL-1993 (first entry)

DE Generic peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 DT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parentally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

Query Match 67.9% Score 38; DB 1; Length 32;  
 Best Local Similarity 70.0%; Pred. No. 1.1; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 3;

QY 1 KFFKFFKFFK 10  
 II II II I  
 DB 2 KFLKFLKFLK 11

RESULT 7  
 R33532 ID R33532 standard; peptide; 11 AA.  
 AC R33532;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 DT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 10; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trip - Leu/Ile/Val)n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parentally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 67.9% Score 38; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 0.43; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 3;

QY 1 KFFKFFKFFK 10  
 II II II I  
 DB 2 KFLKFLKFLK 11

RESULT 8  
 R39294 ID R39294 standard; peptide; 11 AA.  
 AC R39294;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (FORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 DT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 67.9% Score 38; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 0.43; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 3;

QY 1 KFFKFFKFFK 10  
 II II II I  
 DB 2 KFLKFLKFLK 11

RESULT 9  
 R71779 ID R71779 standard; peptide; 10 AA.  
 AC R71779;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.

PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 67.9%; Score 38; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.39;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
 II II II II  
 DB 1 KFLKFLKFLK 10

## RESULT 10

ID R71791 standard; peptide; 12 AA.  
 AC R71791;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide bond 1..12  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 22; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 12 AA;

Query Match 67.9%; Score 38; DB 1; Length 12;  
 Best Local Similarity 70.0%; Pred. No. 0.46;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
 II II II II II  
 DB 2 KFLKFLKFLK 11

## RESULT 11

ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 67.9%; Score 38; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 0.43;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
 II II II II II  
 DB 2 KFLKFLKFLK 11

## RESULT 12

ID W21596 standard; peptide; 10 AA.  
 AC W21596;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #8.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 12; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are



CC correspondingly reduced.  
SQ Sequence 10 AA;

Query Match 67.9%; Score 38; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.39;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
|| || || ||  
Db 1 KFLKFLKFLK 10

## RESULT 13

W21608  
ID W21608 standard; peptide; 12 AA.

AC W21608;

DE 26-AUG-1997 (first entry)

DT Antibiotic potentiating peptide #20.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide bond 1..12

PN W09638183-AI.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M. Varra M;

WPI: 97-034095/03.

DR Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Claim 24; Page 26; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to

CC potentiate the activity of an antibiotic when they are co-administered

CC with the antibiotic. Compositions containing these peptides are used

CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of

CC the outer bacterial membrane to the antibiotic, allowing a reduction

CC in the dose of antibiotic required by 10-50% of the normal dose for

CC in vivo or in vitro application. Any toxic side effects are

CC correspondingly reduced.

SQ Sequence 12 AA;

Query Match 67.9%; Score 38; DB 1; Length 12;  
Best Local Similarity 70.0%; Pred. No. 0.46;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFFKFFKFFK 10  
|| || || ||  
Db 2 KFLKFLKFLK 11

## RESULT 14

W83904  
ID W83904 standard; Protein; 166 AA.

AC W83904;

DT 15-FEB-1999 (first entry)

DE Porcine interferon-gamma.

KW Interferon-gamma; immune interferon; virucide; antiviral;

KW antitumour; pig.

OS Sus scrofa.

FH Key Location/Qualifiers

FT Peptide 1..23

FT /label= Sig\_peptide

FT 24..166

FT /label= Mat\_protein

FT /note= "Claim 1"

PN US5831023-A.

PD 03-NOV-1998.

PF 19-MAY-1995; 444454.

PR 02-OCT-1987; US-104461.  
PR 08-MAR-1982; US-355298.  
PR 01-NOV-1982; US-438128.  
PR 23-AUG-1991; US-749371.  
PR 21-SEP-1992; US-949327.  
PR 19-MAY-1995; US-444454.  
PR (GETH ) GENENTECH INC.  
PI Capon DJ, Goeddel DV;  
DR WPI: 99-007993/01.  
DR N-PSDB; V63230.  
PT Animal interferon polypeptides - useful as antiviral and antitumour  
PT agents  
PS Claim 1; Fig 14C; 48pp; English.  
CC This is the amino acid sequence of porcine interferon gamma, as  
CC deduced from the open reading frame of a cDNA clone (see V63230).  
CC The invention provides bovine, murine, porcine, feline, laprine  
CC and rat mature alpha, beta and gamma interferons and their  
CC propeptides (see W83895-905, W83918 and W70582-85). Recombinant  
CC DNA methods can be utilised to prepare these non-human animal  
CC interferons in amounts sufficient to enable the determination of  
CC their biochemical properties and bioactivities, and hence to allow  
CC efficient production for commercial or biological exploitation.  
CC Interferons are useful as antiviral and antitumour agents.  
SQ Sequence 166 AA;

Query Match 67.9%; Score 38; DB 1; Length 166;  
Best Local Similarity 66.7%; Pred No. 4.8;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFFK 10  
:||||: ||  
Db 76 YFKFFKFFK 84

## RESULT 15

W73233

ID W73233 standard; Protein; 166 AA.

AC W73233;

DT 01-MAR-1999 (first entry)

DE Porcine interferon-gamma.

KW Interferon-gamma; immune interferon; virucide; antiviral;

KW antitumour; pig.

OS Sus scrofa.

FH Key Location/Qualifiers

FT Peptide 1..23

FT /label= Sig\_peptide

FT 24..166

FT /label= Mat\_protein

PN US5827694-A.

PD 27-OCT-1998.

PF 19-MAY-1995; 446171.

PR 02-OCT-1987; US-104461.

PR 08-MAR-1982; US-395298.

PR 01-NOV-1982; US-438128.

PR 23-AUG-1991; US-749371.

PR 21-SEP-1992; US-949327.

PR 19-MAY-1995; US-446171.

PA (GETH ) GENENTECH INC.

PI Capon DJ, Goeddel DV;

DR WPI: 98-593994/50.

DR N-PSDB; V08186.

PT DNA encoding non-human interferon polypeptide(s) - useful for

PT producing recombinant polypeptide(s) to determine their

PT bioactivity

PS Claim 6; Fig 14C; 59pp; English.

CC This is the amino acid sequence of porcine interferon gamma,

CC deduced from a cDNA (see V08186). The invention provides

CC bovine, porcine, feline and rabbit mature interferons and their

CC propeptides (see W73224-35). Recombinant DNA methods can be used

CC to prepare these non-human animal interferons in amounts sufficient

CC to enable the determination of their biochemical properties and

CC bioactivities, and hence to allow efficient production for

CC commercial or biological exploitation. Interferons are useful as  
CC antiviral and antitumor agents. A replicable expression vehicle  
CC comprising claimed non-human animal interferon nucleic acid, a  
CC process for producing such an interferon, and a microorganism  
CC (especially Escherichia coli) or cell culture transfected with the  
CC nucleic acid are claimed.  
SQ Sequence 166 AA;

Query Match 67.9%; Score 38; DB 1; Length 166;  
Best Local Similarity 66.7%; Pred. NO. 4.8;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKKFKFKK 10  
:||||: ||  
Db 76 YKAFKFKK 84

Search completed: September 7, 1999, 20:37:13  
Job time: 18452 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:41 ; Search time 148.39 Seconds  
(without alignments)  
3.733 Million cell updates/sec

Title: US-09-124-280A-11

Perfect score: 39

Sequence: 1 KLLKLLKLL 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL\_10:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	87.2	24	4	014473
2	33	84.6	374	10	080776
3	32	82.1	339	1	058318
4	32	82.1	184	2	065997
5	32	82.1	246	2	051324
6	32	82.1	700	3	013875
7	32	82.1	323	5	018623
8	32	82.1	512	5	076787
9	32	82.1	498	5	076788
10	32	82.1	576	5	076791
11	32	82.1	544	5	076792
12	32	82.1	381	5	076793
13	32	82.1	562	5	076794
14	32	82.1	706	12	089581
15	32	82.1	706	12	089926
16	32	82.1	138	13	091278
17	31	79.5	73	2	050837
18	31	79.5	1032	4	094894
19	31	79.5	646	5	045076
20	31	79.5	1645	5	099140
21	31	79.5	665	5	018597
22	31	79.5	767	5	076977
23	31	79.5	1032	11	062780
24	30	76.9	389	1	027747
25	30	76.9	684	1	058969
26	30	76.9	361	2	067410
27	30	76.9	229	4	095661
28	30	76.9	470	5	023037
29	30	76.9	346	5	018370

30	30	76.9	638	5	019770	019770 caenorhabdi
31	30	76.9	734	5	022145	022145 caenorhabdi
32	30	76.9	884	5	022706	022706 caenorhabdi
33	30	76.9	867	10	092UI6	092UI6 arabidopsis
34	30	76.9	1200	10	092UF2	092UF2 arabidopsis
35	30	76.9	123	10	092TN4	092TN4 petunia hyb
36	30	76.9	1102	11	088877	088877 rattus norv
37	30	76.9	1017	11	089048	089048 rattus norv
38	30	76.9	1087	11	089047	089047 rattus norv
39	30	76.9	235	12	055739	055739 chilo iride
40	30	76.9	119	12	083954	083954 orgyia pseu
41	30	76.9	822	12	056280	056280 orgyia pseu
42	30	76.9	822	12	039989	039989 human herpe
43	30	76.9	822	12	096910	096910 human herpe
44	29	74.4	259	1	028703	028703 archaeoglob
45	29	74.4	506	11	070507	070507 mus musculu

ALIGNMENTS

RESULT 1

014473

ID 014473 PRELIMINARY; PRT; 24 AA.

AC 014473

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DE HYPOTHETICAL 2.9 KD PROTEIN (FRAGMENT).

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RA SVOBODA P., ZAVADIL J.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF020742; AAB71869.1; -

KW Hypothetical protein.

FT NON\_TER 1 1

FT NON\_TER 24 24

SQ SEQUENCE 24 AA; 2921 MW; F58BD8A2 CRC32;

Query Match 87.2%; Score 34; DB 4; Length 24;  
Best Local Similarity 88.9%; Pred. No. 6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KLLKLLKLL 9  
| | | | | | | |  
DB 2 KLLKLLKYL 10

RESULT 2

080776

ID 080776 PRELIMINARY; PRT; 374 AA.

AC 080776

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DE HYPOTHETICAL 2.9 KD PROTEIN (FRAGMENT).

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;

OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;

OC Arabidopsis.

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-CV, COLUMBIA;

RC ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,

RA SIKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,

RA SOMERVILLE C.R., VENTER J.C.;

RT "Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC004481; AAC27403.1; -.  
 SQ SEQUENCE 374 AA; 43653 MW; 45BE65A6 CRC32;

Query Match 84.6%; Score 33; DB 10; Length 374;  
 Best Local Similarity 66.7%; Pred. No. 95;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
 |||:||||  
 Db 77 KLMKLMKLM 85

## RESULT 3

ID O58318 PRELIMINARY; PRT; 339 AA.  
 AC O58318;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE 339AA LONG HYPOTHETICAL PROTEIN.  
 GN PH0598.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE; 98344137.  
 RA KAWABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KOSHIDA N., OGUCHI A.,  
 RA AKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT "Complete sequence and gene organization of the genome of a  
 RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL; AP000002; BAA29687.1; -.  
 SQ SEQUENCE 339 AA; 39403 MW; BB52AF0E CRC32;

Query Match 82.1%; Score 32; DB 1; Length 339;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
 |||:||||  
 Db 291 KLLRLKFL 299

## RESULT 4

ID O65997 PRELIMINARY; PRT; 184 AA.  
 AC O65997;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE SIGX.  
 GN SIGX.  
 OS Clostridium acetobutylicum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Clostridium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 792;  
 RA BEHRNS S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58131; AAC12856.1; -.  
 SQ SEQUENCE 184 AA; 21881 MW; 9BD7EAD2 CRC32;

Query Match 82.1%; Score 32; DB 2; Length 184;

Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KLLKLLKLL 9  
 |||:||||  
 Db 121 KLLKALKIL 129

## RESULT 5

OS1324 PRELIMINARY; PRT; 246 AA.  
 ID O51324;  
 AC O51324;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 29.0 KD PROTEIN.  
 GN BB0349.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN C.,  
 RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL; AE001141; AAC66741.1; -.  
 DR TIGR; BB0349; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 246 AA; 28986 MW; 783FC056 CRC32;

Query Match 82.1%; Score 32; DB 2; Length 246;  
 Best Local Similarity 77.8%; Pred. No. 1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
 |||:||||  
 Db 87 KILKLLKTL 95

## RESULT 6

ID O13875 PRELIMINARY; PRT; 700 AA.  
 AC O13875;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE HYPOTHETICAL 81.0 KD PROTEIN ClB3.10C IN CHROMOSOME I PRECURSOR.  
 GN SPAC1B3.10C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WOOD V.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: SOME, TO C.ELEGANS SEL-1.  
 DR EMBL; Z98598; CAB11247.1; -.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 36 POTENTIAL.  
 FT CHAIN 37 700 HYPOTHETICAL PROTEIN ClB3.10C.  
 FT DOMAIN 573 576 POLY-TYR.

```
SQ SEQUENCE 700 AA; 81048 MW; DF675212 CRC32;

Query Match      82.1%; Score 32; DB 3; Length 700;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 8
   :|:|:|:|:|
Db 656 KLLKLLRL 663

RESULT 7
Q18623 Q18623 PRELIMINARY; PRT; 323 AA.
AC Q18623;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE SIMILAR TO F22B7.7 AND F22B7.8.
GN C44E12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPEX T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA BENTLEY D.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39647; AAA80376.1; -.
SQ SEQUENCE 323 AA; 37315 MW; F730602E CRC32;

Query Match      82.1%; Score 32; DB 5; Length 323;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9
   :|:|:|:|:|
Db 75 RILKLLKL 83

RESULT 8
Q76787 Q76787 PRELIMINARY; PRT; 512 AA.
AC Q76787;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M1-2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;
RT "Analysis of genomic organization of nematode 4TM potassium
RT channels.";
RL Ann. N. Y. Acad. Sci. 0:0-0(1998).
DR EMBL; AF083647; AAC32858.1; -.
SQ SEQUENCE 512 AA; 58455 MW; AB3418EE CRC32;
```

```
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;
RT "Analysis of genomic organization of nematode 4TM potassium
RT channels.";
RL Ann. N. Y. Acad. Sci. 0:0-0(1998).
DR EMBL; AF083643; AAC32854.1; -.
KW Ionic channel.
SQ SEQUENCE 512 AA; 58455 MW; AB3418EE CRC32;

Query Match      82.1%; Score 32; DB 5; Length 512;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9
   :|:|:|:|:|
Db 68 RILKLLKL 76

RESULT 9
Q76788 Q76788 PRELIMINARY; PRT; 498 AA.
AC Q76788;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M1-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;
RT "Analysis of genomic organization of nematode 4TM potassium
RT channels.";
RL Ann. N. Y. Acad. Sci. 0:0-0(1998).
DR EMBL; AF083644; AAC32855.1; -.
KW Ionic channel.
SQ SEQUENCE 498 AA; 56895 MW; 7516ECD5 CRC32;

Query Match      82.1%; Score 32; DB 5; Length 498;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9
   :|:|:|:|:|
Db 68 RILKLLKL 76

RESULT 10
Q76791 Q76791 PRELIMINARY; PRT; 576 AA.
AC Q76791;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M2-2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;
RT "Analysis of genomic organization of nematode 4TM potassium
RT channels.";
RL Ann. N. Y. Acad. Sci. 0:0-0(1998).
DR EMBL; AF083647; AAC32858.1; -.
KW Ionic channel.
SQ SEQUENCE 576 AA; 65550 MW; EC4A6BE9 CRC32;
```

Query Match 82.1%; Score 32; DB 5; Length 576;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 Db 132 RILKLLKIL 140

RESULT 11  
 076792 PRELIMINARY; PRT; 544 AA.  
 AC 076792;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M2-3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;  
 RT "Analysis of genomic organization of nematode 4TM potassium  
 channels.";  
 RL Ann. N. Y. Acad. Sci. 0:0-0(1998).  
 DR EMBL; AF083648; AAC32859.1; -.  
 KW Ionic channel.  
 SQ SEQUENCE 544 AA; 62141 MW; 3995682C CRC32;

Query Match 82.1%; Score 32; DB 5; Length 544;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 Db 132 RILKLLKIL 140

RESULT 12  
 076793 PRELIMINARY; PRT; 381 AA.  
 AC 076793;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M3 (FRAGMENT).  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;  
 RT "Analysis of genomic organization of nematode 4TM potassium  
 channels.";  
 RL Ann. N. Y. Acad. Sci. 0:0-0(1998).  
 DR EMBL; AF083649; AAC32860.1; -.  
 KW Ionic channel.  
 FT NON\_TER 381  
 SQ SEQUENCE 381 AA; 43205 MW; 9BC61399 CRC32;

Query Match 82.1%; Score 32; DB 5; Length 381;  
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 Db 75 RILKLLKIL 83

RESULT 13  
 076794 PRELIMINARY; PRT; 562 AA.  
 AC 076794;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE PUTATIVE POTASSIUM CHANNEL SUBUNIT N2P17M2-1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WANG Z.-W., KUNKEL M.T., WEI A., BUTLER A., SALKOFF L.;  
 RT "Analysis of genomic organization of nematode 4TM potassium  
 channels.";  
 RL Ann. N. Y. Acad. Sci. 0:0-0(1998).  
 DR EMBL; AF083651; AAC32862.1; -.  
 KW Ionic channel.  
 SQ SEQUENCE 562 AA; 63990 MW; 8089ACE8 CRC32;

Query Match 82.1%; Score 32; DB 5; Length 562;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 Db 132 RILKLLKIL 140

RESULT 14  
 089581 PRELIMINARY; PRT; 706 AA.  
 AC 089581;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE HELICASE.  
 GN Q706L.  
 OS African swine fever virus (ASFV).  
 OC Viruses; dsDNA viruses, no RNA stage;  
 OC African swine fever-like viruses.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE; 96036500.  
 RA YANEZ R.J., RODRIGUEZ J.M., NOGAL M.L., YUSTE L., ENRIQUEZ C.,  
 RODRIGUEZ J.F., VINUELA E.;  
 RT "Immune protection conferred by the baculovirus-related glycoprotein  
 of Thogoto virus (Orthomyxoviridae).";  
 RL Virology 208:249-278(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE; 94233765.  
 RA LA VEGA I., GONZALEZ A., BLASCO R., CALVO V., VINUELA E.;  
 RT "Nucleotide sequence and variability of the inverted terminal  
 repetitions of African swine fever virus DNA.";  
 RL Virology 201:152-156(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE; 90219205.  
 RA GONZALEZ A., CALVO V., ALMAZAN F., ALMENDRAL J.M., RAMIREZ J.C.,  
 LA VEGA I., BLASCO R., VINUELA E.;  
 RT "Multigene families in African swine fever virus: family 360.";  
 RL J. Virol. 64:2073-2081(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE; 90219204.

RA ALMENDRAL J.M., ALMAZAN F., BLASCO R., VINUELA E.;  
 RT "Multigene families in African swine fever virus: family 110.";  
 RL J. Virol. 64:2064-2072(1990).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 91134988.  
 RA CAMACHO A., VINUELA E.;  
 RT "Protein p22 of African swine fever virus: an early structural  
 protein that is incorporated into the membrane of infected cells.";  
 RL Virology 181:251-257(1991).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 94187118.  
 RA ALMAZAN F., MURGUIA J.R., RODRIGUEZ J.M., LA VEGA I., VINUELA E.;  
 RT Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 94187118.  
 RA RODRIGUEZ J.M., YANEZ R.J., PAN R., RODRIGUEZ J.F., SALAS M.L.,  
 VINUELA E.;  
 RT "Multigene families in African swine fever virus: family 505.";  
 RL J. Virol. 68:2746-2751(1994).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93346971.  
 RA YANEZ R.J., RODRIGUEZ J.M., RODRIGUEZ J.F., SALAS M.L., VINUELA E.;  
 RT "African swine fever virus thymidylate kinase gene: sequence and  
 transcriptional mapping.";  
 RL J. Gen. Virol. 74:1633-1638(1993).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 94065656.  
 RA ALCAMI A., ANGULO A., VINUELA E.;  
 RT Mapping and sequence of the gene encoding the African swine fever  
 virion protein of M(r) 11500.";  
 RL J. Gen. Virol. 74:2317-2324(1993).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93277388.  
 RA MUNOZ M., FREIJE J.M., SALAS M.L., VINUELA E., LOPEZ-OTIN C.;  
 RT "Structure and expression in E. coli of the gene coding for protein  
 p10 of African swine fever virus.";  
 RL Arch. Virol. 130:93-107(1993).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 90357780.  
 RA BLASCO R., LOPEZ-OTIN C., MUNOZ M., BOCKAMP E.O., SIMON-MATEO C.,  
 VINUELA E.;  
 RT "Sequence and evolutionary relationships of African swine fever virus  
 thymidine kinase.";  
 RL Virology 178:301-304(1990).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93281390.  
 RA YANEZ R.J., BOURSNEILL M., NOGAL M.L., YUSTE L., VINUELA E.;  
 RT "African swine fever virus encodes two genes which share significant  
 homology with the two largest subunits of DNA-dependent RNA  
 polymerases.";  
 RL Nucleic Acids Res. 21:2423-2427(1993).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93353606.  
 RA RODRIGUEZ J.M., YANEZ R.J., ALMAZAN F., VINUELA E., RODRIGUEZ J.F.;  
 RT "African swine fever virus encodes a CD2 homolog responsible for the  
 adhesion of erythrocytes to infected cells.";

J. Virol. 67:5312-5320(1993).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 94085774.  
 RA YANEZ R.J., RODRIGUEZ J.M., BOURSNEILL M., RODRIGUEZ J.F., VINUELA E.;  
 RT "Two putative African swine fever virus helicases similar to yeast  
 'DEAH' pre-mRNA processing proteins and vaccinia virus ATPases Dill  
 and D6R.";  
 RL Gene 134:161-174(1993).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 90233993.  
 RA LOPEZ-OTIN C., FREIJE J.M., PARRA F., MENDEZ E., VINUELA E.;  
 RT "Mapping and sequence of the gene coding for protein p72, the major  
 capsid protein of African swine fever virus.";  
 RL Virology 175:477-484(1990).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 94123986.  
 RA RODRIGUEZ J.M., YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;  
 RT "The DNA polymerase-encoding gene of African swine fever virus:  
 sequence and transcriptional analysis.";  
 RL Gene 136:103-110(1993).  
 RN [17]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93327788.  
 RA SIMON-MATEO C., ANDRES G., VINUELA E.;  
 RT "Polyprotein processing in African swine fever virus: a novel gene  
 expression strategy for a DNA virus.";  
 RL EMBO J. 12:2977-2987(1993).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93233210.  
 RA PRADOS F.J., VINUELA E., ALCAMI A.;  
 RT "Sequence and characterization of the major early phosphoprotein p32  
 of African swine fever virus.";  
 RL J. Virol. 67:2475-2485(1993).  
 RN [19]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 92260660.  
 RA ALCAMI A., ANGULO A., LOPEZ-OTIN C., MUNOZ M., FREIJE J.M.,  
 CARRASCOA A.L., VINUELA E.;  
 RT "Amino acid sequence and structural properties of protein p12, an  
 African swine fever virus attachment protein.";  
 RL J. Virol. 66:3860-3868(1992).  
 RN [20]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93174976.  
 RA YANEZ R.J., VINUELA E.;  
 RT "African swine fever virus encodes a DNA ligase.";  
 RL Virology 193:531-536(1993).  
 RN [21]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 93174941.  
 RA PENA L., YANEZ R.J., REVILLA Y., VINUELA E., SALAS M.L.;  
 RT "African swine fever virus guanylyltransferase.";  
 RL Virology 193:319-328(1993).  
 RN [22]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE: 95159428.  
 RA SIMON-MATEO C., FREIJE J.M., ANDRES G., LOPEZ-OTIN C., VINUELA E.;  
 RT "Mapping and sequence of the gene encoding protein p17, a major  
 African swine fever virus structural protein.";  
 RL Virology 206:1140-1144(1995).

```

RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE; 92263807.
RA GARCIA-BEATO R., FREIJE J.M., LOPEZ-OTIN C., BLASCO R., VINUELA E.,
RA SALAS M.L.;
RT "A gene homologous to topoisomerase II in African swine fever
RL virus.";
RL Virology 188:938-947(1992).
RN [24]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE; 94091056.
RA FREIJE J.M., LAIN S., VINUELA E., LOPEZ-OTIN C.;
RT "Nucleotide sequence of a nucleoside triphosphate phosphohydrolase
Query Match 82.1%; Score 32; DB 12; Length 706;
Best Local Similarity 87.5%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKLLKLL 9
DB 114 LKLIKLL 121

Search completed: September 7, 1999, 20:34:42
Job time: 19738 sec

RESULT 15
O89926 PRELIMINARY; PRT; 706 AA.
AC Q89926;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE HELICASE.
GN J10L.
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage;
OC African swine fever-like viruses.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RX MEDLINE; 94014996.
RA VYDELINGUM S., BAYLIS S.A., BRISTOW C., SMITH G.L., DIXON L.K.;
RT "Duplicated genes within the variable right end of the genome of a
RT pathogenic isolate of African swine fever virus.";
RL J. Gen. Virol. 74:2125-2130(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL20 /1;
RX MEDLINE; 94292916.
RA DIXON L.K., TWIGG S.R.F., BAYLIS S.A., VYDELINGUM S., BRISTOW C.,
RA HAMMOND J.M., SMITH G.L.;
RT "Nucleotide sequence of a 55 kbp region from the right end of the
RT genome of a pathogenic African swine fever virus isolate (Malawi
RL LIL20/1).";
RL J. Gen. Virol. 7:1655-1684(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93393435.
RA ROBERTS P.C., LU Z., KUTISH G.F., ROCK D.L.;
RT "Three adjacent genes of African swine fever virus with similarity to
RT essential poxvirus genes.";
RL Arch. Virol. 132:331-342(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI;
RX MEDLINE; 93389456.
RA BAYLIS S.A., TWIGG S.R.F., VYDELINGUM S., DIXON L.K., SMITH G.L.;
RT "Three African swine fever virus genes encoding proteins with
RT homology to putative helicases of vaccinia virus.";
RL J. Gen. Virol. 74:1969-1974(1993).
DR EMBL; X71982; CAA50829.1; -
DR EMBL; M88275; AAA03222.1; -
DR EMBL; X72952; CAA51457.1; -

```

```

DR PFAM; PF00271; helicase_C; 1.
KW Helicase.
SQ SEQUENCE 706 AA; 80498 MW; 1FA08FD0 CRC32;

```

```

Query Match 82.1%; Score 32; DB 12; Length 706;
Best Local Similarity 87.5%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 LKLLKLL 9
DB 114 LKLIKLL 121

```

```

Search completed: September 7, 1999, 20:34:42
Job time: 19738 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:13 ; Search time 71.87 Seconds  
(without alignments)  
3.540 Million cell updates/sec

Title: US-09-124-280A-11

Perfect score: 39

Sequence: 1 KLLKLLKLL 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	84.6	246	1 RL7_DICDI	P11874 dictyostell
2	33	84.6	339	1 RMAR_CANGA	P21358 candida gla
3	33	84.6	2493	1 YB4_YEAST	P35194 saccharomyc
4	32	82.1	917	1 YGJ_YEAST	P53148 saccharomyc
5	31	79.5	811	1 REXL_YEAST	P48743 saccharomyc
6	30	76.9	440	1 CHLB_CYPAR	P48099 cyanophora
7	30	76.9	884	1 SECA_OILIU	Q32743 olithodisc
8	30	76.9	560	1 TATR_NPVCF	P41716 choristoneu
9	30	76.9	560	1 TATR_NPVOP	P22114 orgyia pseu
10	30	76.9	822	1 VGLB_HSV7J	P52352 herpes simp
11	29	74.4	879	1 MYSP_DROME	P35415 drosophila
12	29	74.4	2413	1 PROB_YEAST	P33334 saccharomyc
13	29	74.4	2165	1 RRPL_HRSVA	P28887 human respi
14	29	74.4	885	1 YDQ3_YEAST	Q06639 saccharomyc
15	29	74.4	21	1 YD90_HAEIN	P45194 haemophilus
16	29	74.4	822	1 YJC9_YEAST	P47061 saccharomyc
17	28	71.8	956	1 CB31_YEAST	P32504 saccharomyc
18	28	71.8	757	1 CC5_SCHPO	P39964 schizosacch
19	28	71.8	1852	1 CIGL_CIPCA	P22316 cyprinus ca
20	28	71.8	676	1 LSHR_CALJA	O02721 callithrix
21	28	71.8	699	1 LSHR_HUMAN	P22888 homo sapien
22	28	71.8	1419	1 MDR_PLAFF	P13568 plasmodium
23	28	71.8	404	1 NANH_CLOSO	P15698 clostridium
24	28	71.8	409	1 PEPT_SALTY	P26311 salmonella
25	28	71.8	433	1 PRES_STAAU	P03864 staphylococ
26	28	71.8	496	1 REPR_STRAG	P18629 streptococc
27	28	71.8	496	1 REPS_STRPY	P14752 streptococc
28	28	71.8	250	1 RL7B_SCHPO	P25457 schizosacch
29	28	71.8	261	1 RNG6_MOUSE	P28078 mus musculu
30	28	71.8	878	1 SECA_ANTSP	Q06461 antithannio
31	28	71.8	782	1 TALA_POVM3	P03074 mouse polyo
32	28	71.8	785	1 TALA_POVMC	P03073 mouse polyo
33	28	71.8	777	1 TALA_POVMC	P12905 mouse polyo
34	28	71.8	440	1 TAMI_POVM3	P03076 mouse polyo
35	28	71.8	421	1 TAMI_POVMA	P03077 mouse polyo
36	28	71.8	421	1 TAMI_POVMC	P12906 mouse polyo
37	28	71.8	195	1 TASM_POVMA	P03078 mouse polyo
38	28	71.8	2077	1 TEGU_HSV6G	P30002 herpes simp
39	28	71.8	2077	1 TEGU_HSV6U	P52340 herpes simp
40	28	71.8	198	1 Y264_MYCGE	P47506 mycoplasma
41	28	71.8	1374	1 YAH8_SCHPO	Q09884 schizosacch
42	28	71.8	563	1 YEBK_SCHPO	Q13816 schizosacch
43	28	71.8	1420	1 YM8B_YEAST	Q03496 saccharomyc

44 27 69.2 541 1 AAAT\_RABIT 019105 oryctolagus  
45 27 69.2 633 1 YR45\_CAEEL 009362 caenorhabdi

ALIGNMENTS

RESULT 1

RL7\_DICDI STANDARD; PRT; 246 AA.  
AC P11874;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE 60S RIBOSOMAL PROTEIN L7.  
GN RPL7.  
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).  
OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AX2;  
RX MEDLINE; 89345108.  
RA SZYMKOWSKI D.E., KELLY B., DEERING R.A.;  
RT "A dictyostelium discoideum cDNA coding for a protein with homology  
to the rat ribosomal protein L7".  
RL NUCLEIC ACIDS RES. 17:5393-5393(1989).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE; 90370447.  
RA SZYMKOWSKI D.E., DEERING R.A.;  
RT "Identification and characterization of a Dictyostelium discoideum  
ribosomal protein gene."  
RL NUCLEIC ACIDS RES. 18:4695-4701(1990).  
CC -!- FUNCTION: BINDS TO G-RICH STRUCTURES IN 28S RRNA AND IN MRNAS.  
CC PLAYS A REGULATORY ROLE IN THE TRANSLATION APPARATUS; INHIBITS  
CC CELL-FREE TRANSLATION OF MRNAS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
EMBL: X14909; G7357; -  
PIR: S04849; R5D07.  
DR DICTYDB; D001011; -  
DR PROSITE; PS00634; RIBOSOMAL\_L30; 1.  
DR PFAM; PF00327; L30; 1.  
KW RIBOSOMAL PROTEIN; RNA-BINDING.  
SQ SEQUENCE 246 AA; 27932 MW; 72367CC8 CRC32;

Query Match 84.6%; Score 33; DB 1; Length 246;

Best Local Similarity 77.8%; Pred. NO. 14;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9

1:|||||  
Db 105 KVLKLLRLL 113

RESULT 2

RMAR\_CANGA STANDARD; PRT; 339 AA.  
ID RMAR\_CANGA  
AC P21358;  
DT 01-MAY-1991 (REL. 18, CREATED)  
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL RIBOSOMAL PROTEIN VARI.  
GN VARI.

OS CANDIDA GLABRATA (YEAST) (TORULOPSIS GLABRATA).  
 OG MITOCHONDRION.  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETEALES;  
 OC CANDIDACEAE; CANDIDA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86011564.  
 RA AINLEY W.M., MACREADIE I.G., BUTOW R.A.;  
 RT "VARI Gene on the mitochondrial genome of *Torulopsis glabrata*.";  
 RL J. MOL. BIOL. 184:565-576(1985).  
 CC -!- FUNCTION: ESSENTIAL FOR MITOCHONDRIAL PROTEIN SYNTHESIS AND  
 CC REQUIRED FOR THE MATURATION OF SMALL RIBOSOMAL SUBUNITS  
 CC (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -!- SIMILARITY: BELONGS TO THE VARI FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X02893; G600448; -.  
 DR PIR; S04682; S04682.  
 KW RIBOSOMAL PROTEIN; MITOCHONDRION.  
 SQ SEQUENCE 339 AA; 40828 MW; B93BDD51 CRC32;

Query Match 84.6%; Score 33; DB 1; Length 339;  
 Best Local Similarity 87.5%; Pred. No. 19;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 8  
 ||:|||||  
 Db 2 KLMKLLKL 9

RESULT 3  
 YB44\_YEAST  
 ID YB44\_YEAST STANDARD; PRT; 2493 AA.  
 AC P35194;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 287.5 KD PROTEIN IN PDR3-HTA2 INTERGENIC REGION.  
 GN YBL004W OR YBL0101.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETEALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE OF 1-2066 FROM N.A.  
 RC STRAIN-S288C;  
 RA RIEGER M.;  
 RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE OF 1214-2493 FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 94378721.  
 RA WOLFE K.H., LOHAN A.J.E.;  
 RT "Sequence around the centromere of *Saccharomyces cerevisiae*  
 RT chromosome II: similarity of CEN2 to CEN4.";  
 RL YEAST 10:S41-S46(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z35765; G535982; -.  
 DR EMBL; Z26494; G403312; -.  
 DR PIR; S45734; S45734.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 1170 1187 POTENTIAL.  
 SQ SEQUENCE 2493 AA; 287487 MW; B66AB060 CRC32;

Query Match 84.6%; Score 33; DB 1; Length 2493;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9  
 ||:|||||  
 Db 1242 KILKILKL 1250

RESULT 4  
 YGJ3\_YEAST  
 ID YGJ3\_YEAST STANDARD; PRT; 917 AA.  
 AC P53148;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 104.8 KD PROTEIN IN PAN2-NUP145 INTERGENIC REGION.  
 GN YGL093W.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETEALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 97435481.  
 RA RIEGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;  
 RT "Sequence analysis of 203 kilobases from *Saccharomyces cerevisiae*  
 RT chromosome VII.";  
 RL YEAST 13:1077-1090(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z72615; E243976; -.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 917 AA; 104825 MW; 9389FB90 CRC32;

Query Match 82.1%; Score 32; DB 1; Length 917;  
 Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9  
 ||:|||||  
 Db 817 KVILKLLKL 825

RESULT 5  
 RFXL\_YEAST  
 ID RFXL\_YEAST STANDARD; PRT; 811 AA.  
 AC P48743;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 90.6 KD PROTEIN IN CBFS-DKAI INTERGENIC REGION.  
 GN YLR176C OR L9470.18.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETEALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU 2.,  
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,  
RA KUCABA T., HALLSWORTH K., HAWKINS J., HILLIER L., JIER M.,  
RA JOHNSON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,  
RA MARDIS E., MENEZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,  
RA RIFKEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,  
RA WILCOX L., WOHLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;  
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U17446; G2340986; -;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 811 AA; 90583 MW; 792768F2 CRC32;  
  
Query Match 79.5%; Score 31; DB 1; Length 811;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 KLLKLLKLL 9  
Db 650 KLIKLLKFI 658  
  
RESULT 6  
CHLB\_CYAPA  
ID CHLB\_CYAPA STANDARD; PRT; 440 AA.  
AC P49099;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE PROTOCHLOROPHYLLIDE REDUCTASE CHLB SUBUNIT (EC 1.3.1.33) (NADPH-  
DE PROTOCHLOROPHYLLIDE OXIDOREDUCTASE CHLB SUBUNIT).  
GN CHLB.  
OS CYANOPHORA PARADOXA.  
OG CYANELLE.  
OC EURARYOTA; GLAUCOCYSTOPHYCEAE; CYANOPHORACEAE; CYANOPHORA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STREIN-LB555 / PRINGSHEIM;  
RA STREWALT V.L., MICHALOWSKI C.B., LUFFELHARDT W., BOHNERT H.J.,  
RA BRYANT D.A.;  
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: INVOLVED IN THE LIGHT-INDEPENDENT ACCUMULATION OF  
CC CHLOROPHYLL, PROBABLY AT THE STEP OF REDUCTION OF PROTO-  
CC CHLOROPHYLLIDE TO CHLOROPHYLL (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: CHLOROPHYLLIDE A + NADP(+) -  
CC PROTOCHLOROPHYLLIDE + NADPH.  
CC -!- PATHWAY: LIGHT-INDEPENDENT CHLOROPHYLL BIOSYNTHESIS.  
CC -!- SIMILARITY: BELONGS TO THE CHLB/BCHB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U30821; G1016093; -;  
DR PFAM; PF00148; oxidored.nitro; 2.  
DR MENDEL; 7871; CYAPA.chlb;1.  
KW CYANELLE; PHOTOSYNTHESIS; CHLOROPHYLL BIOSYNTHESIS; OXIDOREDUCTASE;  
NADP.

SQ SEQUENCE 440 AA; 50284 MW; 9E06A9F7 CRC32;  
  
Query Match 76.9%; Score 30; DB 1; Length 440;  
Best Local Similarity 87.5%; Pred. No. 89;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 KLLKLLKLL 8  
Db 277 KLLKLLNL 284  
  
RESULT 7  
SECA\_OLILU  
ID SECA\_OLILU STANDARD; PRT; 884 AA.  
AC Q32743;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PREPROTEIN TRANSLOCASE SECA SUBUNIT.  
GN SECA.  
OS OLISTHODISCUS LUTEUS (HETEROSIGMA AKASHIWO).  
OG CHLOROPLAST.  
OC EURARYOTA; STRAMENOPILES; RAPHIIDOPHYCEAE; OLISTHODISCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA VALENTIN K.U., FISCHER S., VOGEL H.;  
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: INVOLVED IN PROTEIN EXPORT AND IMPORT. SECA HAS A  
CC CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF ATP TO THE TRANSFER OF  
CC PRE-SECRETORY PROTEINS ACROSS THE MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -!- SIMILARITY: BELONGS TO THE SECA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z35718; G527574; -;  
DR PROSITE; PS01312; SECA; 1.  
DR PFAM; PF01043; SecA\_protein; 1.  
DR MENDEL; 4234; OLILU;seca;1.  
KW PROTEIN TRANSPORT; ATP-BINDING; CHLOROPLAST; TRANSLOCATION; TRANSPORT.  
FT NP\_BIND 97 104 ATP (BY SIMILARITY).  
SQ SEQUENCE 884 AA; 102799 MW; 068E8C6E CRC32;  
  
Query Match 76.9%; Score 30; DB 1; Length 884;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 KLLKLLKLL 9  
Db 501 KIIRLLKIL 509  
  
RESULT 8  
TATR\_NPVCF  
ID TATR\_NPVCF STANDARD; PRT; 560 AA.  
AC P41716;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE TRANS-ACTIVATING TRANSCRIPTIONAL REGULATORY PROTEIN (IMMEDIATE EARLY  
DE PROTEIN 1) (IE-1).  
GN IE1.  
OS CHORISTONEURA FUMIFERANA NUCLEAR POLYHEDROSIS VIRUS (CFMNPV).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;  
OC NUCLEOPOLYHEDROVIRUS.

CC SEQUENCE FROM N.A.  
RA KUZIO J., SCHODELLA E., FAULKNER P.;  
RL SUBMITTED (XXX-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: REGULATORY TRANSCRIPTIONAL PROTEIN, WHICH TRANS-  
CC ACTIVATES GENE EXPRESSION FROM EARLY BACULOVIRUS PROMOTERS. CAN  
CC ALSO TRANS-ACTIVATE ITS OWN PROMOTER, SUGGESTING THAT IT IS  
CC AUTOREGULATED DURING NORMAL INFECTION OF INSECT CELLS.  
CC -!- SIMILARITY: TO OTHER BACULOVIRUSES IE-1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: L04945; G332513; -  
KW EARLY PROTEIN; TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR.  
SQ SEQUENCE 560 AA; 64946 MW; 2F8A9B6 CRC32;  
  
Query Match 76.9%; Score 30; DB 1; Length 560;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 LKLLKLL 9  
DB 527 LKLLQLL 534  
||||:||  
-----  
RESULT 9  
TATR\_NPVOP STANDARD; PRT; 560 AA.  
ID TATR\_NPVOP AC P22114;  
DT 01-AUG-1991 (REL. 19, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE TRANS-ACTIVATING TRANSCRIPTIONAL REGULATORY PROTEIN (IMMEDIATE EARLY  
DE PROTEIN 1) (IE-1).  
GN IE1.  
OS ORGYIA PSEUDOTSUGATA MULTICAPSID POLYHEDROSIS VIRUS (OPNPV).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;  
OC NUCLEOPOLYHEDROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 9111968.  
RA THEILMANN D.A., STEWART S.;  
RT "Identification and characterization of the IE-1 gene of Orgyia  
RT pseudotsugata multicapsid nuclear polyhedrosis virus.";  
RL VIROLOGY 180:492-508(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97271300.  
RA AHRENS C.H., RUSSELL R.R., FUNK C.J., EVANS J., HARWOOD S.,  
RA ROHRMANN G.F.;  
RT "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear  
RT polyhedrosis virus genome";  
RL VIROLOGY 229:381-399(1997).  
CC -!- FUNCTION: REGULATORY TRANSCRIPTIONAL PROTEIN, WHICH TRANS-  
CC ACTIVATES GENE EXPRESSION FROM EARLY BACULOVIRUS PROMOTERS. CAN  
CC ALSO TRANS-ACTIVATE ITS OWN PROMOTER, SUGGESTING THAT IT IS  
CC AUTOREGULATED DURING NORMAL INFECTION OF INSECT CELLS.  
CC -!- ALTERNATIVE PRODUCTS: A LARGER TRANSCRIPT MAY ARISE BY ALTERNATIVE  
CC SPLICING OF THE SAME GENE.  
CC -!- SIMILARITY: TO OTHER BACULOVIRUSES IE-1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M63414; G332528; -  
DR EMBL: U75930; G3911391; -  
DR PIR: A38544; RGNVPM.  
KW EARLY PROTEIN; TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR;  
KW ALTERNATIVE SPLICING.  
FT DOMAIN 1 132 ASP/GLU-RICH (ACIDIC).  
SQ SEQUENCE 560 AA; 64315 MW; 7517CB5 CRC32;  
  
Query Match 76.9%; Score 30; DB 1; Length 560;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 LKLLKLL 9  
DB 527 LKLLQLL 534  
||||:||  
-----  
RESULT 10  
VGLB\_HSV7J STANDARD; PRT; 822 AA.  
ID VGLB\_HSV7J AC P52352;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE GLYCOPROTEIN B PRECURSOR.  
DE GB OR U39.  
OS HERPES SIMPLEX VIRUS (TYPE 7 / STRAIN JI) (HHV7).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;  
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA NICHOLAS J.;  
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN B.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: U43400; G1139641; -  
DR FRAM; PF00606; Glycoprotein B; 1.  
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 822 GLYCOPROTEIN B.  
FT DOMAIN 23 649 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 650 670 POTENTIAL.  
FT DOMAIN 671 683 POTENTIAL.  
FT TRANSMEM 684 704 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 705 822 POTENTIAL.  
FT CARBOHYD 152 152 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 244 244 POTENTIAL.  
FT CARBOHYD 324 324 POTENTIAL.  
FT CARBOHYD 338 338 POTENTIAL.  
FT CARBOHYD 352 352 POTENTIAL.  
FT CARBOHYD 358 358 POTENTIAL.  
FT CARBOHYD 387 387 POTENTIAL.  
FT CARBOHYD 390 390 POTENTIAL.  
FT CARBOHYD 482 482 POTENTIAL.  
FT CARBOHYD 517 517 POTENTIAL.  
FT CARBOHYD 567 567 POTENTIAL.  
SQ SEQUENCE 822 AA; 93147 MW; 60FA195A CRC32;  
  
Query Match 76.9%; Score 30; DB 1; Length 822;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
1:11 :111  
777 KMLKAIKLL 785

Db

RESULT 11  
MYSP\_DROME STANDARD; PRT; 879 AA.  
AC P35415;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PARAMYOSIN, LONG FORM.  
GN PRM.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CANTON-S; TISSUE-MUSCLE;  
RX MEDLINE: 92113004.  
RA BECKER K.D., O'DONNELL P.T., HEITZ J.M., VITO M., BERNSTEIN S.I.;  
RT "Analysis of Drosophila paramyosin: identification of a novel isoform  
which is restricted to a subset of adult muscles.";  
RL J. CELL BIOL. 116:669-681(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 9217957.  
RA VINOS J., MAROTO M., GARESSE R., MARCO R., CERVERA M.;  
RT "Drosophila melanogaster paramyosin: developmental pattern, mapping  
and properties deduced from its complete coding sequence.";  
RL MOL. GEN. GENET. 231:385-394(1992).  
RN [3]  
RP SEQUENCE OF 103-571 FROM N.A., AND CHARACTERIZATION.  
RC STRAIN-OREGON-R;  
RX MEDLINE: 91332901.  
RA VINOS J., DOMINGO A., MARCO R., CERVERA M.;  
RT "Identification and characterization of Drosophila melanogaster  
paramyosin.";  
RL J. MOL. BIOL. 220:687-700(1991).  
CC -!- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF  
MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.  
CC -!- SUBUNIT: HETERODIMER OF TWO ISOFORMS (POTENTIAL).  
CC -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL LARVAL AND ADULT MUSCLE  
TISSUES. EXPRESSION IS FIVE TIMES HIGHER IN TUBULAR THAN IN  
FIBRILLAR MUSCLES.  
CC -!- DEVELOPMENTAL STAGE: UNDETECTABLE DURING GASTRULATION AND EARLY  
PHASES OF GERM BAND FORMATION. INCREASES DURING ORGANOGENESIS,  
AROUND 10 HOURS POSTFERTILIZATION, TO THE ADULT STAGE.  
CC -!- PTM: THE MORE ACIDIC AND LESS-ABUNDANT ISOFORM IS PHOSPHORYLATED  
IN VIVO.  
CC -!- ALTERNATIVE PRODUCTS: THERE ARE AT LEAST THREE ISOFORMS THAT ARISE  
FROM ALTERNATIVE SPLICING AND/OR POSTTRANSLATIONAL MODIFICATIONS.  
CC -!- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X62590; G11031; -  
DR EMBL: X58722; G10959; -  
DR PIR: S20486; S20486.  
DR PIR: S22028; S22028.  
DR FLYBASE: FBgn0003149; Prm.  
DR HSP: P02649; 1LE4.

KW COILED COIL; HEPTAD REPEAT PATTERN; MUSCLE PROTEIN; THICK FILAMENT;  
KW MYOSIN; PHOSPHORYLATION; ALTERNATIVE SPLICING.  
FT DOMAIN 1 31 NON-HELICAL REGION (POTENTIAL).  
FT DOMAIN 32 858 COILED COIL (POTENTIAL).  
FT DOMAIN 859 879 NON-HELICAL REGION (POTENTIAL).  
FT DISULFID 368 368 INTERCHAIN (POTENTIAL).  
FT DISULFID 784 784 INTERCHAIN (POTENTIAL).  
FT CONFLICT 500 500 MISSING (IN REF. 2).  
SQ SEQUENCE 879 AA; 102338 MW; 46DFE4D5 CRC32;

Query Match 74.4%; Score 29; DB 1; Length 879;  
Best Local Similarity 77.8%; Pred. No. 2.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
:1111 :111  
Db 99 ELLKRLKLL 107

RESULT 12  
PRO8\_YEAST STANDARD; PRT; 2413 AA.  
AC P33334;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE PRE-MRNA SPLICING FACTOR PRP8.  
GN PRP8 OR RNA8 OR DBF3 OR DNA39 OR YHR165C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 95304847.  
RA HOGES P.E., JACKSON S.P., BROWN J.D., BEGGS J.D.;  
RT "Extraordinary sequence conservation of the PRP8 splicing factor.";  
RL YEAST 11:337-342(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 95140615.  
RA SHEA J.E., TOIN J.H., JOHNSTON L.H.;  
RT "The budding yeast U5 snRNP Prp8 is a highly conserved protein which  
links RNA splicing with cell cycle progression.";  
RL NUCLEIC ACIDS RES. 22:5555-5564(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RX MEDLINE: 94378003.  
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,  
DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,  
KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,  
LATREILLE P., LOUIS E.J., MACRI C., MARIS E., MENEZES S., MOUSER L.,  
NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,  
VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,  
VAUDIN M.;  
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
VIII.";  
RL SCIENCE 265:2077-2082(1994).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE: 88216580.  
RA JACKSON S.P., LOSSKY M., BEGGS J.D.;  
RT "Cloning of the RNA8 gene of Saccharomyces cerevisiae, detection of  
the RNA8 protein, and demonstration that it is essential for nuclear  
pre-mRNA splicing.";  
RL MOL. CELL. BIOL. 8:1067-1075(1988).  
CC -!- FUNCTION: INVOLVED IN PRE-MRNA SPLICING. U5 SNRNP PROTEIN.  
CC APPEARS TO CONTACT THE PRE-MRNA DURING SPLICING. ALSO HAS A ROLE  
IN CELL CYCLE. BINDS RNA.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- SIMILARITY: STRONG, TO C.ELEGANS PROTEIN C50C3.6.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; 224732; G395157; -;  
 DR EMBL; L29421; G460042; -;  
 DR EMBL; U00027; G551320; -;  
 DR PIR; S34670; S34670; -;  
 DR PIR; S48905; S48905; -;  
 DR SGD; L0001500; PRP8.  
 KW MRNA PROCESSING; MRNA SPLICING; SPLICEOSOME; NUCLEAR PROTEIN;  
 KW RNA-BINDING.  
 FT DOMAIN 5 9 POLY-PRO.  
 FT DOMAIN 20 27 POLY-PRO.  
 FT DOMAIN 50 56 POLY-PRO.  
 FT DOMAIN 72 78 POLY-PRO.  
 FT CONFLICT 388 420  
 FT PHLYNSRPRSVRIWPNFVSCIIQNDEYDTP ->  
 FT LIVIIPGVQCAVHGIIICRVLSRTMRSTRL  
 FT (IN REF. 2).  
 FT T -> S (IN REF. 2).  
 FT W -> C (IN REF. 2).  
 FT CONFLICT 1132 1132  
 FT CONFLICT 1575 1575  
 FT SEQUENCE 2413 AA; 279501 MW; 340180AD CRC32;

Query Match 74.4%; Score 29; DB 1; Length 2413;  
 Best Local Similarity 55.6%; Pred. No. 7.3e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 :|:|:|:|:  
 DB 624 ELKKMKLI 632

RESULT 13  
 RRPL\_HRSVA  
 ID RRPL\_HRSVA STANDARD; PRT; 2165 AA.  
 AC P28887;  
 DT 01-DEC-1992 (REL. 24, CREATED)  
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE RNA POLYMERASE BETA SUBUNIT (EC 2.7.7.48) (LARGE STRUCTURAL PROTEIN)  
 DE (L PROTEIN).  
 GN L.  
 OS HUMAN RESPIRATORY SYNCYTIAL VIRUS (STRAIN A2).  
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;  
 OC PARAMYXOVIRIDAE; PNEUMOVIRINAE; PNEUMOVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91272488.  
 RA STEC D.S., HILL M.G. III, COLLINS P.L.;  
 RT "Sequence analysis of the polymerase L gene of human respiratory  
 RT syncytial virus and predicted phylogeny of nonsegmented  
 RT negative-strand viruses.";  
 RL VIROLOGY 183:273-287(1991).  
 CC -!- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY  
 CC SYNTHESIS OF MRNA SYNTHESIS, CAPPING, METHYLATION AND POLY(A)  
 CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAS, RNA EDITING OF THE P  
 CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.  
 CC -!- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M75730; G333956; -;

DR PIR; A40317; RRNZAA2.  
 DR PFAM; PF00946; Paramyx\_RNA\_pol; 1.  
 KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE.  
 SQ SEQUENCE 2165 AA; 250384 MW; 72CE0E98 CRC32;

Query Match 74.4%; Score 29; DB 1; Length 2165;  
 Best Local Similarity 75.0%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 8  
 :|:|:|:|:  
 DB 403 KFLKLIKL 410

RESULT 14  
 YD03\_YEAST  
 ID YD03\_YEAST STANDARD; PRT; 885 AA.  
 AC Q06639;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE PUTATIVE 101.7 KD TRANSCRIPTIONAL REGULATORY PROTEIN IN PROL-CPRS  
 DE INTERGENIC REGION.  
 DE YDR303C OR D9740.13.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,  
 RA KUCABA T., HALLSMORTH K., HAWKINS J., HILLIER L., JIER M.,  
 RA JOHNSON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,  
 RA MARDIS E., MENEZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,  
 RA RIEFEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,  
 RA WILCOX L., WOHLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DDJ DATA BANKS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CC CLUSTER DOMAIN.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U28374; G849220; -;  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_1; 1.  
 DR PROSITE; PS50048; ZN2\_CYS6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; ZN2\_Cys; 1.  
 DR HSSP; P08657; 1CLD.  
 KW HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING;  
 KW NUCLEAR PROTEIN; ZINC; METAL-BINDING  
 FT DNA\_BIND 14 42  
 FT ZN(2)-CYS(6), FUNGAL-TYPE.  
 SQ SEQUENCE 885 AA; 101720 MW; 8FA6C9BF CRC32;

Query Match 74.4%; Score 29; DB 1; Length 885;  
 Best Local Similarity 55.6%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 :|:|:|:|:  
 DB 602 KIIQLIKLL 610

RESULT 15  
 YD90\_HAEN

```
ID YD90 HAEIN STANDARD; PRT; 21 AA.
AC P45194;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN H11390.
GN H11390
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20;
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
RA UTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDER D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32819; G1574229; -
DR TIGR; H11390; -
DR PROSITE; PS01097; HUPF_HYPC; UNKNOWN_1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 21 AA; 2387 MW; 85BB62A0 CRC32;
-----
Query Match 74.4%; Score 29; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 6.8;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 KLLKLLKLL 9
DB 8 KLSKLMKIL 16
|||:|:|
|||:|:|
```

Search completed: September 7, 1999, 23:50:15  
Job time: 1954 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:25 ; Search time 116.8 Seconds  
(without alignments)  
3.087 Million cell updates/sec

Title: US-09-124-280A-11  
Perfect score: 39  
Sequence: 1 KLLKLLKLL 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	246	1 R5D07	ribosomal protein
2	33	84.6	374	2 T02315	hypothetical prote
3	33	84.6	2493	2 S45734	probable membrane
4	32	82.1	706	2 S42174	NTPase - African s
5	32	82.1	706	2 J02210	probable nucleic a
6	32	82.1	246	2 D70143	hypothetical prote
7	32	82.1	917	2 S64100	probable membrane
8	32	82.1	339	2 B71175	hypothetical prote
9	32	82.1	138	2 S36115	interferon - Japan
10	31	79.5	1032	2 A57514	RNA helicase HEL1
11	31	79.5	73	2 C70255	hypothetical prote
12	31	79.5	771	2 S51421	hypothetical prote
13	31	79.5	1645	2 A37792	spectrin beta-H ch
14	30	76.9	560	1 RGNVPM	trans-activating t
15	30	76.9	389	2 B09096	corrinoid/iron-sul
16	30	76.9	884	2 S49216	preprotein translo
17	30	76.9	684	2 E64496	eIF-4A family prob
18	30	76.9	235	2 T03076	hypothetical prote
19	30	76.9	361	2 H70422	hypothetical prote
20	29	74.4	2165	1 RRN2A2	genome polyprotein
21	29	74.4	415	1 B250	streptokinase (BC
22	29	74.4	274	2 E71730	2,3,4,5-tetrahydro
23	29	74.4	1175	2 S26874	DNA-directed RNA p
24	29	74.4	878	2 S20486	paramyosin - fruit
25	29	74.4	879	2 S22028	paramyosin, standa
26	29	74.4	885	2 S51189	probable membrane
27	29	74.4	369	2 B70220	conserved hypothet
28	29	74.4	301	2 F70121	hypothetical prote
29	29	74.4	21	2 F64121	hypothetical prote
30	29	74.4	425	2 D71518	probable polyva pol
31	29	74.4	429	2 C59379	chemotaxis histidi
32	29	74.4	259	2 H69445	conserved hypothet
33	29	74.4	292	2 F69372	osmoprotection pro
34	29	74.4	385	2 S69587	hypothetical prote
35	29	74.4	822	2 S56801	hypothetical prote
36	29	74.4	2413	2 S34670	splicing factor PR
37	29	74.4	853	2 S49876	gamma-adaptin - sm
38	29	74.4	1922	2 T00637	hypothetical prote
39	29	74.4	208	2 S72286	ribosomal protein

ALIGNMENTS

RESULT 1  
R5D07

ribosomal protein L7 - slime mold (Dictyostellium discoideum)

C:Species: Dictyostellium discoideum

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 05-Sep-1997

C:Accession: S04849; S12834

R:Szymkowski, D.E.; Kelly, B.; Deering, R.A.

Nucleic Acids Res. 17, 5393, 1989

A:Title: A Dictyostellium discoideum cDNA coding for a protein with homology to the ra

A:Reference number: S04849; MUID:89345108

A:Accession: S04849

A:Molecule type: mRNA

A:Residues: 1-246 <SZY>

A:Cross-references: EMBL:X14909; NID:g7356; PID:g7357

C:Genetics:

A:Gene: rpl7

C:Superfamily: rat ribosomal protein L7

C:Keywords: protein biosynthesis; ribosome

Query Match 84.6%; Score 33; DB 1; Length 246;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
Db 105 KVLKLLRL 113

RESULT 2  
T02315

hypothetical protein Fl3p17.11 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 05-Mar-1999

C:Accession: T02315

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC Fl3p17 genomic sequence.

A:Reference number: Z14168

A:Accession: T02315

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-374 <ROU>

A:Cross-references: EMBL:AC004481; NID:g3337347; PID:g3337358

C:Genetics:

A:Map position: 2

A>Note: Fl3p17.11

Query Match 84.6%; Score 33; DB 2; Length 374;  
Best Local Similarity 66.7%; Pred. No. 31;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
Db 77 KLMKLMKLM 85

RESULT 3  
S45734

probable membrane protein YBL004w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBL0101

C:Species: Saccharomyces cerevisiae  
C:Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 21-Nov-1997  
C:Accession: S45734; S45733; S44556; S37318  
R:Rieger, M.  
submitted to the Protein Sequence Database, August 1994  
A:Reference number: S45734  
A:Accession: S45734  
A:Molecule type: DNA  
A:Residues: 1-2066 <RIE>  
A:Cross-references: EMBL:Z35765; MIPS:YBL004W  
A:Experimental source: strain S288C  
R:Lohan, A.J.E.; Wolfe, K.H.  
submitted to the Protein Sequence Database, August 1994  
A:Reference number: S45730  
A:Accession: S45733  
A:Molecule type: DNA  
A:Residues: 1214-2493 <LOH>  
A:Cross-references: EMBL:Z35765; MIPS:YBL004W  
A:Experimental source: strain S288C  
R:Wolfe, K.H.; Lohan, A.J.E.  
Yeast 10(Suppl.A), S41-S46, 1994  
A:Title: Sequence around the centromere of Saccharomyces cerevisiae chromosome II: similar  
A:Reference number: S44556  
A:Accession: S44556  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1214-2493 <WOL>  
A:Cross-references: EMBL:Z26494; NID:g403311; PID:g403312  
A:Experimental source: strain S288C  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993  
C:Genetics:  
A:Map position: 2L  
C:Keywords: transmembrane protein  
F:165-182/Domain: transmembrane #status predicted <TM1>  
F:350-366/Domain: transmembrane #status predicted <TM2>  
F:656-672/Domain: transmembrane #status predicted <TM3>  
F:1021-1037/Domain: transmembrane #status predicted <TM4>  
F:1170-1187/Domain: transmembrane #status predicted <TM5>  
F:1195-1211/Domain: transmembrane #status predicted <TM6>  
F:1606-1622/Domain: transmembrane #status predicted <TM7>  
F:1656-1674/Domain: transmembrane #status predicted <TM8>  
F:1763-1779/Domain: transmembrane #status predicted <TM9>  
F:1910-1926/Domain: transmembrane #status predicted <TM10>  
F:2093-2109/Domain: transmembrane #status predicted <TM11>  
F:2355-2371/Domain: transmembrane #status predicted <TM12>

Query Match 84.6%; Score 33; DB 2; Length 2493;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
I:|||||  
Db 1242 KILKILKLI 1250

RESULT 4  
NTPase - African swine fever virus  
C:Species: African swine fever virus, ASFV  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Sep-1997  
C:Accession: S42174  
R:Freije, J.M.P.; Lain, S.; Vinuela, E.; Lopez-Otin, C.  
Virus Res. 30, 63-72, 1993  
A:Title: Nucleotide sequence of a nucleotide triphosphate phosphohydrolase gene from Afr  
A:Reference number: S42174  
A:Accession: S42174  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-706 <FRE>  
A:Cross-references: EMBL:X69952; NID:g407358; PID:g407359

Query Match 82.1%; Score 32; DB 2; Length 706;  
Best Local Similarity 87.5%; Pred. No. 88;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLKLLKLL 9  
I:|||||  
Db 114 LLKLLKLL 121

RESULT 5  
Probable nucleic acid-dependent ATPase - African swine fever virus  
N:Alternate names: helicase homolog j10L protein  
C:Species: African swine fever virus, ASFV  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Sep-1997  
C:Accession: JQ2210; S27894  
R:Baylis, S.A.; Twigg, S.R.F.; Vydelingum, S.; Dixon, L.K.; Smith, G.L.  
J. Gen. Virol. 74, 1969-1974, 1993  
A:Title: Three African swine fever virus genes encoding proteins with homology to put  
A:Reference number: JQ2209  
A:Accession: JQ2210  
A:Molecule type: DNA  
A:Residues: 1-706 <BAY>  
A:Cross-references: EMBL:X72952; NID:g414092; PID:g414093  
R:Roberts, P.C.; Lu, Z.; Rock, D.L.  
submitted to the EMBL Data Library, July 1992  
A:Description: Three adjacent genes of African swine fever virus with homologies to e  
A:Reference number: S27892  
A:Accession: S27894  
A:Molecule type: DNA  
A:Residues: 1-706 <ROB>  
A:Cross-references: EMBL:M88275; NID:g210602; PID:g210605  
C:Genetics:  
A:Gene: j10L

Query Match 82.1%; Score 32; DB 2; Length 706;  
Best Local Similarity 87.5%; Pred. No. 88;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLKLLKLL 9  
I:|||||  
Db 114 LLKLLKLL 121

RESULT 6  
Hypothetical protein BB0349 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
C:Accession: D70143  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943  
A:Accession: D70143  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-246 <KLE>  
A:Cross-references: GB:AE001141; GB:AE000783; NID:g2688250; PID:g2688263; TIGR:BB0349  
A:Experimental source: strain B31

Query Match 82.1%; Score 32; DB 2; Length 246;  
Best Local Similarity 77.8%; Pred. No. 31;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
I:|||||  
Db 87 KILKLLKTL 95

```
RESULT 7
S64100
probable membrane protein YGL093w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G3168
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 14-Nov-1997
C:Accession: S64100
R:Kieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64100
A:Molecule type: DNA
A:Residues: 1-917 <RIE>
A:Cross-references: EMBL:272615; NID:g1322624; PID:e243976; PID:g1322625; MIPS:YGL093w
A:Experimental source: strain S288C
C:Genetics:
C:Keywords: transmembrane protein
F:851-867/Domain: transmembrane #status predicted <TMM>

Query Match 82.1%; Score 32; DB 2; Length 917;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9
| : |||||
Db 817 KVIKLLKLL 825

RESULT 8
B71175
hypothetical protein PH0598 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 14-Aug-1998
C:Accession: B71175
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137
A:Accession: B71175
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-339 <KAW>
A:Cross-references: GB:AP000002; NID:g3236129; PID:d1030630; PID:g3257004
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0598

Query Match 82.1%; Score 32; DB 2; Length 339;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9
| : |||||
Db 291 KLLRLKLL 299

RESULT 9
S36115
interferon - Japanese flounder
C:Species: Paralichthys olivaceus (Japanese flounder)
C:Date: 09-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 12-Feb-1999
C:Accession: S36115
R:Tamai, T.; Shirahata, S.; Noguichi, T.; Sato, N.; Kimura, S.; Murakami, H.
Biochim. Biophys. Acta 1174, 182-186, 1993
A:Title: Cloning and expression of flatfish (Paralichthys olivaceus) interferon cDNA.
A:Reference number: S36115; MUID:93363636
```

```
A:Accession: S36115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-138 <TAM>

Query Match 82.1%; Score 32; DB 2; Length 138;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLKLLKLL 9
| : |||||
Db 41 LLLKLLKLI 48

RESULT 10
A57514
RNA helicase HEL117 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 05-Jun-1998
C:Accession: A57514
R:Sukeyawa, J.; Blobel, G.
J. Biol. Chem. 270, 15702-15706, 1995
A:Title: A putative mammalian RNA helicase with an arginine-serine-rich domain colocal
A:Reference number: A57514; MUID:95318159
A:Accession: A57514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1032 <SUK>
A:Cross-references: GB:U25746; NID:g897914; PID:g897915
C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology
C:Keywords: ATP; P-loop
F:416-712/Domain: DEAD/H box helicase homology <DEAD>
F:416-423/Region: nucleotide-binding motif A (P-loop)
F:525-530/Region: nucleotide-binding motif B
F:529-532/Region: DEAD motif

Query Match 79.5%; Score 31; DB 2; Length 1032;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9
| : |||||
Db 605 KFLKLELL 613

RESULT 11
C70255
hypothetical protein BBK34 - Lyme disease spirochete plasmid K/lp36
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Jun-1998
C:Accession: C70255
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: C70255
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-73 <KLE>
A:Cross-references: GB:AE000788; NID:g2690123; PID:g2690158; TIGR:BBK34
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 79.5%; Score 31; DB 2; Length 73;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 KLLKLLKLL 9  
||:|:|:|:  
Db 64 KLVKLMKVL 72

RESULT 12  
S51421  
hypothetical protein YLR176c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein L9470.18  
C:Species: Saccharomyces cerevisiae  
C>Date: 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 06-Feb-1998  
C:Accession: S51421  
R:Wohldmann, P.  
submitted to the EMBL Data Library, November 1994  
A:Description: The sequence of S. cerevisiae cosmid 9470.  
A:Reference number: S51421  
A:Accession: S51421  
A:Molecule type: DNA  
A:Residues: 1-771 <WOH>  
A:Cross-references: EMBL:U17246; NID:g577192; PID:g577210; MIPS:YLR176c  
C:Genetics:  
A:Gene: SGD:REF1  
A:Cross-references: SGD:S0004166; MIPS:YLR176c  
A:Map position: 12R

Query Match 79.5%; Score 31; DB 2; Length 771;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
||:|:|:|:  
Db 610 KLIKLLKFI 618

RESULT 13  
A37792  
spectrin beta-H chain - fruit fly (Drosophila melanogaster) (fragment)  
C:Species: Drosophila melanogaster  
C>Date: 30-Apr-1991 #sequence\_revision 08-Nov-1996 #text\_change 12-Feb-1999  
C:Accession: A37792; S70848; S15666  
R:Dubreuil, R.R.; Byers, T.J.; Stewart, C.T.; Kiehart, D.P.  
J. Cell Biol. 111, 1849-1858, 1990  
A:Title: A beta-spectrin isoform from Drosophila (beta-H) is similar in size to vertebrate  
A:Reference number: A37792; MUID:91035599  
A:Accession: A37792  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1645 <DUB>  
A:Cross-references: GB:X53992  
A:Note: Met-14 is the probable initiator  
R:Dubreuil, R.R.  
submitted to the EMBL Data Library, July 1990  
A:Reference number: S70848  
A:Accession: S70848  
A:Molecule type: mRNA  
A:Residues: 1-400,403-410,'OL',411-1645 <DUW>  
A:Cross-references: EMBL:X53992; NID:g7654; PID:g7655  
A:Note: Met-14 is the probable initiator  
C:Genetics:  
A:Gene: FlyBase:kst  
A:Cross-references: FlyBase:FBgn0004167  
C:Superfamily: alpha-actinin actin-binding domain homology; SH3 homology; spectrin/dystrophin  
C:Keywords: actin binding  
F:36-258/Domain: alpha-actinin actin-binding domain homology <ACT>  
F:288-399/Domain: spectrin/dystrophin repeat homology <SN1>  
F:400-510/Domain: spectrin/dystrophin repeat homology <SN2>  
F:851-896/Domain: SH3 homology <SH3>  
F:990-1091/Domain: spectrin/dystrophin repeat homology <SP2>  
F:1303-1408/Domain: spectrin/dystrophin repeat homology <SP5>  
F:1409-1512/Domain: spectrin/dystrophin repeat homology <SP1>

---

Query Match 79.5%; Score 31; DB 2; Length 1645;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
||:|:|:|:  
Db 70 KLIKLLLEII 78

RESULT 14  
RGNVPK  
trans-activating transcription regulator - Orgyia pseudotsugata multicapsid nuclear P  
N:Alternate names: IE-1 protein  
C:Species: Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus, OpMPV  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 02-Jul-1998  
C:Accession: A38544  
R:Theilmann, D.A.; Stewart, S.  
Virology 180, 492-508, 1991  
A:Title: Identification and characterization of the IE-1 gene of Orgyia pseudotsugata  
A:Reference number: A38544; MUID:91111968  
A:Accession: A38544  
A:Molecule type: DNA  
A:Residues: 1-560 <THE>  
A:Cross-references: EMBL:M63414; NID:g332527; PID:g332528  
C:Superfamily: Autographa californica nuclear polyhedrosis virus trans-activating tra  
C:Keywords: alternative splicing; early protein; transcription

Query Match 76.9%; Score 30; DB 1; Length 560;  
Best Local Similarity 87.5%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKLLKLL 9  
||:|:|:|:  
Db 527 LKLLQLL 534

RESULT 15  
B69096  
corrinoid/iron-sulfur protein, small subunit - Methanobacterium thermoautotrophicum (C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
C:Accession: B69096  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69090; MUID:98037514  
A:Accession: B69096  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-389 <MTH>  
A:Cross-references: GB:AE000928; GB:AE000666; NID:g2622835; PID:g2622844  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1712  
A:Start codon: GTG  
C:Superfamily: corrinoid/iron-sulfur protein small chain

Query Match 76.9%; Score 30; DB 2; Length 389;  
Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
||:|:|:|:  
Db 5 KTELKLL 13

Search completed: September 7, 1999, 23:06:27

Job time: 2475 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:35 ; Search time 80.79 Seconds  
(without alignments)  
1.099 Million cell updates/sec

Title: US-09-124-280A-11  
Perfect score: 39  
Sequence: 1 KLLKLLKLL 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	100.0	9	1	US-08-097-830E-11	Sequence 11, Appl
2	39	100.0	9	2	US-08-456-112B-11	Sequence 11, Appl
3	33	84.6	17	1	US-07-725-331-49	Sequence 49, Appl
4	33	84.6	17	1	US-07-725-331-51	Sequence 51, Appl
5	33	84.6	17	1	US-07-725-331-53	Sequence 53, Appl
6	33	84.6	22	1	US-07-725-331-60	Sequence 60, Appl
7	33	84.6	26	1	US-07-725-331-61	Sequence 61, Appl
8	33	84.6	30	1	US-07-725-331-62	Sequence 62, Appl
9	33	84.6	36	1	US-07-725-331-63	Sequence 63, Appl
10	33	84.6	11	1	US-08-193-521-1	Sequence 1, Appl
11	33	84.6	12	1	US-08-193-521-2	Sequence 2, Appl
12	33	84.6	13	1	US-08-193-521-3	Sequence 3, Appl
13	33	84.6	14	1	US-08-193-521-4	Sequence 4, Appl
14	33	84.6	16	1	US-08-193-521-6	Sequence 6, Appl
15	33	84.6	11	1	US-08-434-120-95	Sequence 95, Appl
16	33	84.6	12	1	US-08-434-120-96	Sequence 96, Appl
17	33	84.6	13	1	US-08-434-120-97	Sequence 97, Appl
18	33	84.6	14	1	US-08-434-120-98	Sequence 98, Appl
19	33	84.6	16	1	US-08-434-120-100	Sequence 100, Appl
20	33	84.6	11	1	US-08-465-325-94	Sequence 94, Appl
21	33	84.6	12	1	US-08-465-325-95	Sequence 95, Appl
22	33	84.6	13	1	US-08-465-325-96	Sequence 96, Appl
23	33	84.6	14	1	US-08-465-325-97	Sequence 97, Appl
24	33	84.6	16	1	US-08-465-325-99	Sequence 99, Appl
25	33	84.6	16	2	US-08-569-188-1	Sequence 1, Appl
26	33	84.6	16	2	US-08-569-188-2	Sequence 2, Appl
27	33	84.6	17	2	US-08-569-188-3	Sequence 3, Appl
28	33	84.6	17	2	US-08-569-188-4	Sequence 4, Appl
29	33	84.6	18	2	US-08-569-188-5	Sequence 5, Appl
30	33	84.6	18	2	US-08-569-188-6	Sequence 6, Appl
31	33	84.6	14	2	US-08-569-188-8	Sequence 8, Appl
32	33	84.6	16	2	US-08-569-188-10	Sequence 10, Appl
33	33	84.6	16	2	US-08-569-188-11	Sequence 11, Appl
34	33	84.6	16	2	US-08-569-188-12	Sequence 12, Appl
35	33	84.6	16	2	US-08-569-188-13	Sequence 13, Appl
36	33	84.6	17	2	US-08-569-188-14	Sequence 14, Appl
37	33	84.6	18	2	US-08-569-188-15	Sequence 15, Appl
38	33	84.6	17	3	PCT-US91-05047-49	Sequence 49, Appl
39	33	84.6	17	3	PCT-US91-05047-51	Sequence 51, Appl

40 33 84.6 17 3 PCT-US91-05047-53 Sequence 53, Appl  
41 33 84.6 22 3 PCT-US91-05047-60 Sequence 60, Appl  
42 33 84.6 26 3 PCT-US91-05047-61 Sequence 61, Appl  
43 33 84.6 30 3 PCT-US91-05047-62 Sequence 62, Appl  
44 33 84.6 36 3 PCT-US91-05047-63 Sequence 63, Appl  
45 33 84.6 18 3 PCT-US94-07019-15 Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-11  
; Sequence 11, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porcio, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-11

Query Match 100.0%; Score 39; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
Db 1 KLLKLLKLL 9

RESULT 2  
US-08-456-112B-11  
; Sequence 11, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porcio, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8998  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-11

Query Match 100.0%; Score 39; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KLLKLLKLL 9  
Db 1 KLLKLLKLL 9

RESULT 3  
US-07-725-331-49  
Sequence 49, Application US/07725331  
Patent No. 5294605  
GENERAL INFORMATION:  
APPLICANT: Houghten, Richard  
APPLICANT: Blondelle, Sylvie  
TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
TITLE OF INVENTION: Analogues Thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
ADDRESSEE: & Milnanow  
STREET: 180 No. 5294605th Stetson  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/725,331  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,422  
FILING DATE: 19-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.

REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: 421250-80  
TELEPHONE: 3126165418  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: C-terminal amide, acetylated  
OTHER INFORMATION: at N-terminus.  
US-07-725-331-49

Query Match 84.6%; Score 33; DB 1; Length 17;  
Best Local Similarity 88.9%; Pred No. 2.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 KLLKLLKLL 9  
Db 2 KLLKLLKLL 10

RESULT 4  
US-07-725-331-51  
Sequence 51, Application US/07725331  
Patent No. 5294605  
GENERAL INFORMATION:  
APPLICANT: Houghten, Richard  
APPLICANT: Blondelle, Sylvie  
TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
TITLE OF INVENTION: Analogues Thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
ADDRESSEE: & Milnanow  
STREET: 180 No. 5294605th Stetson  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/725,331  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,422  
FILING DATE: 19-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: 421250-80  
TELEPHONE: 3126165418  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: C-terminal amide, acetylated



OTHER INFORMATION: at N-terminus.  
US-07-725-331-51

Query Match 84.6%; Score 33; DB 1; Length 17;  
Best Local Similarity 88.9%; Pred. No. 2.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
Db 6 KLLKLLKLL 14

## RESULT 5

US-07-725-331-53  
; Sequence 53, Application US/07725331  
; Patent No. 5294605  
; GENERAL INFORMATION:  
; APPLICANT: Houghten, Richard  
; APPLICANT: Blondelle, Sylvie  
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
; TITLE OF INVENTION: Analogues Thereof  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
; ADDRESSEE: & Milnamow  
; STREET: 180 No. 5294605th Stetson  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/725.331  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/554,422  
; FILING DATE: 19-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER: 421250-80  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 3126165418  
; TELEFAX: 3126165460  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; OTHER INFORMATION: C-terminal amide, acetylated  
; OTHER INFORMATION: at N-terminus.  
US-07-725-331-53

Query Match 84.6%; Score 33; DB 1; Length 17;  
Best Local Similarity 88.9%; Pred. No. 2.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
Db 6 KLLKLLKLL 14

## RESULT 6

US-07-725-331-60  
; Sequence 60, Application US/07725331  
; Patent No. 5294605  
; GENERAL INFORMATION:  
; APPLICANT: Houghten, Richard  
; APPLICANT: Blondelle, Sylvie  
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
; TITLE OF INVENTION: Analogues Thereof  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
; ADDRESSEE: & Milnamow  
; STREET: 180 No. 5294605th Stetson  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/725.331  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/554,422  
; FILING DATE: 19-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER: 421250-80  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 3126165418  
; TELEFAX: 3126165460  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; OTHER INFORMATION: acetylated at N-terminus, may be a  
; OTHER INFORMATION: C-terminal amide  
US-07-725-331-60

Query Match 84.6%; Score 33; DB 1; Length 22;  
Best Local Similarity 88.9%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
Db 3 KLLKLLKLL 11

## RESULT 7

US-07-725-331-61  
; Sequence 61, Application US/07725331  
; Patent No. 5294605  
; GENERAL INFORMATION:  
; APPLICANT: Houghten, Richard  
; APPLICANT: Blondelle, Sylvie  
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
; TITLE OF INVENTION: Analogues Thereof  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
; ADDRESSEE: & Milnamow  
; STREET: 180 No. 5294605th Stetson  
; CITY: Chicago

STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/725,331  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,422  
FILING DATE: 19-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: 421250-80  
TELEPHONE: 3126165418  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: acetylated at N-terminus, may be  
a C-terminal amide.  
US-07-725-331-61

Query Match 84.6%; Score 33; DB 1; Length 26;  
Best Local Similarity 88.9%; Pred. No. 3.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0;  
Gaps 0;  
OY 1 KLLKLLKLL 9  
DB 7 KLLKLLKLL 15

RESULT 8  
US-07-725-331-62  
Sequence 62, Application US/07725331  
Patent No. 5294605  
GENERAL INFORMATION:  
APPLICANT: Houghten, Richard  
APPLICANT: Blondelle, Sylvie  
TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
Analogues Thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
STREET: 180 No. 5294605th Stetson  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/725,331  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,422

FILING DATE: 19-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: 421250-80  
TELEPHONE: 3126165418  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: acetylated at N-terminus, may be  
a C-terminal amide.  
US-07-725-331-62

Query Match 84.6%; Score 33; DB 1; Length 30;  
Best Local Similarity 88.9%; Pred. No. 3.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0;  
Gaps 0;  
OY 1 KLLKLLKLL 9  
DB 11 KLLKLLKLL 19

RESULT 9  
US-07-725-331-63  
Sequence 63, Application US/07725331  
Patent No. 5294605  
GENERAL INFORMATION:  
APPLICANT: Houghten, Richard  
APPLICANT: Blondelle, Sylvie  
TITLE OF INVENTION: Amphiphilic Peptide Compositions and  
Analogues Thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,  
STREET: 180 No. 5294605th Stetson  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/725,331  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554,422  
FILING DATE: 19-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: 421250-80  
TELEPHONE: 3126165418  
TELEFAX: 3126165460  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
FEATURE: acetylated at N-terminus, may be  
OTHER INFORMATION: a C-terminal amide  
US-07-725-331-63

Query Match 84.6%; Score 33; DB 1; Length 36;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9  
Db 17 KLLKLLKL 25

RESULT 10  
US-08-193-521-1  
Sequence 1, Application US/08193521  
Patent No. 5470950  
GENERAL INFORMATION:  
APPLICANT: Maloy, W. Lee  
APPLICANT: Karl, U. Prasad  
APPLICANT: Williams, Jon I.  
TITLE OF INVENTION: Biologically Active Peptide  
TITLE OF INVENTION: Compositions and Uses Therefor  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,521  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/870,960  
FILING DATE:  
APPLICATION NUMBER: 07/760,054  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 421250-161  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: May be a C-terminal amide,  
and/or may be acetylated at  
OTHER INFORMATION: N-terminus.

Query Match 84.6%; Score 33; DB 1; Length 11;  
Best Local Similarity 88.9%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9  
Db 3 KLLKLLKL 11

RESULT 11  
US-08-193-521-2  
Sequence 2, Application US/08193521  
Patent No. 5470950  
GENERAL INFORMATION:  
APPLICANT: Maloy, W. Lee  
APPLICANT: Karl, U. Prasad  
APPLICANT: Williams, Jon I.  
TITLE OF INVENTION: Biologically Active Peptide  
TITLE OF INVENTION: Compositions and Uses Therefor  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,521  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/870,960  
FILING DATE:  
APPLICATION NUMBER: 07/760,054  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 421250-161  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: May be a C-terminal amide,  
and/or may be acetylated at  
OTHER INFORMATION: N-terminus.

Query Match 84.6%; Score 33; DB 1; Length 12;  
Best Local Similarity 88.9%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKL 9  
Db 4 KLLKLLKL 12

RESULT 12  
US-08-193-521-3  
Sequence 3, Application US/08193521  
Patent No. 5470950

GENERAL INFORMATION:  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
APPLICANT: Williams, Jon I.  
TITLE OF INVENTION: Biologically Active Peptide  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi & Stewart  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,521  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/870,960  
FILING DATE:  
APPLICATION NUMBER: 07/760,054  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 421250-161  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: May be a C-terminal amide,  
and/or may be acetylated at  
OTHER INFORMATION: N-terminus.  
US-08-193-521-3

Query Match 84.6%; Score 33; DB 1; Length 13;  
Best Local Similarity 88.9%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLKLLKLL 9  
DB 5 KLKLLKLL 13

RESULT 13  
US-08-193-521-4  
Sequence 4, Application US/08193521  
Patent No. 5470950  
GENERAL INFORMATION:  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
APPLICANT: Williams, Jon I.  
TITLE OF INVENTION: Biologically Active Peptide  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi & Stewart

STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,521  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/870,960  
FILING DATE:  
APPLICATION NUMBER: 07/760,054  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 421250-161  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
OTHER INFORMATION: May be a C-terminal amide,  
and/or may be acetylated at  
OTHER INFORMATION: N-terminus.  
US-08-193-521-4

Query Match 84.6%; Score 33; DB 1; Length 14;  
Best Local Similarity 88.9%; Pred. No. 1.8;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLKLLKLL 9  
DB 6 KLKLLKLL 14

RESULT 14  
US-08-193-521-6  
Sequence 6, Application US/08193521  
Patent No. 5470950  
GENERAL INFORMATION:  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
APPLICANT: Williams, Jon I.  
TITLE OF INVENTION: Biologically Active Peptide  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi & Stewart  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/193,521  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/870,960  
;; FILING DATE:  
;; APPLICATION NUMBER: 07/760,054  
;; FILING DATE: 13-SEP-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Olstein, Elliot M.  
;; REGISTRATION NUMBER: 24,025  
;; REFERENCE/DOCKET NUMBER: 421250-161  
;; TELEPHONE: 201-994-1700  
;; TELEFAX: 201-994-1744  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 16 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; OTHER INFORMATION: May be a C-terminal amide.  
;; OTHER INFORMATION: and/or may be acetylated at  
;; OTHER INFORMATION: N-terminus.  
US-08-193-521-6

Query Match 84.6%; Score 33; DB 1; Length 16;  
Best Local Similarity 88.9%; Pred. No. 2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
Db 8 KLKLLKLL 16

RESULT 15  
US-08-434-120-95  
; Sequence 95, Application US/08434120  
; Patent No. 5635479  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Margaret A.  
; APPLICANT: Jacob, Leonard S.  
; APPLICANT: Maloy, W. Lee  
; TITLE OF INVENTION: Treatment of Gynecological  
; TITLE OF INVENTION: Malignancies with  
; TITLE OF INVENTION: Biologically Active Peptides  
; NUMBER OF SEQUENCES: 117  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DW4.V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,120  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,950  
; FILING DATE:  
; APPLICATION NUMBER: US/08/226,108  
; FILING DATE:

;; APPLICATION NUMBER: US/07/937,462  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Olstein, Elliot M.  
;; REGISTRATION NUMBER: 24,025  
;; REFERENCE/DOCKET NUMBER: 421250-194  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201-994-1700  
;; TELEFAX: 201-994-1744  
;; INFORMATION FOR SEQ ID NO: 95:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-434-120-95

Query Match 84.6%; Score 33; DB 1; Length 11;  
Best Local Similarity 88.9%; Pred. No. 1.4;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLLKLLKLL 9  
Db 3 KLKLLKLL 11

Search completed: September 7, 1999, 22:38:35  
Job time: 7919 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: September 7, 1999, 20:37:13 ; Search time 147.16 Seconds  
(without alignments)  
1.449 Million cell updates/sec

Title: US-09-124-280A-11  
Perfect score: 39  
Sequence: 1 KLLKLLKLL 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	9	1 R71782	Peptide neutralisi
2	39	100.0	9	1 W21599	Antibiotic potenti
3	34	87.2	37	1 W77378	Lytic peptide with
4	34	84.6	28	1 P91334	Amino acid sequenc
5	33	84.6	28	1 P91335	Amino acid sequenc
6	33	84.6	28	1 P91337	Amino acid sequenc
7	33	84.6	28	1 P91350	Amino acid sequenc
8	33	84.6	17	1 R21386	Sequence of amphip
9	33	84.6	17	1 R21388	Sequence of amphip
10	33	84.6	17	1 R21390	Sequence of amphip
11	33	84.6	22	1 R21397	Sequence of amphip
12	33	84.6	26	1 R21398	Sequence of amphip
13	33	84.6	30	1 R21399	Sequence of amphip
14	33	84.6	36	1 R21400	Sequence of amphip
15	33	84.6	17	1 R22870	Amphiphilic peptid
16	33	84.6	17	1 R22872	Amphiphilic peptid
17	33	84.6	17	1 R22874	Amphiphilic peptid
18	33	84.6	22	1 R22881	Amphiphilic peptid
19	33	84.6	26	1 R22882	Amphiphilic peptid
20	33	84.6	30	1 R22883	Amphiphilic peptid
21	33	84.6	36	1 R22884	Amphiphilic peptid
22	33	84.6	16	1 R31151	C-terminal subst.
23	33	84.6	11	1 R31146	C-terminal subst.
24	33	84.6	12	1 R31147	C-terminal subst.
25	33	84.6	13	1 R31148	C-terminal subst.
26	33	84.6	14	1 R31149	C-terminal subst.
27	33	84.6	11	1 R35364	Amphiphilic peptid
28	33	84.6	12	1 R35365	Amphiphilic peptid
29	33	84.6	13	1 R35366	Amphiphilic peptid
30	33	84.6	14	1 R35367	Amphiphilic peptid
31	33	84.6	16	1 R35369	Amphiphilic peptid
32	33	84.6	12	1 R33957	Amphiphilic peptid
33	33	84.6	13	1 R33958	Amphiphilic peptid
34	33	84.6	14	1 R33959	Amphiphilic peptid
35	33	84.6	11	1 R33956	Amphiphilic peptid
36	33	84.6	16	1 R33961	Amphiphilic peptid
37	33	84.6	11	1 R36371	Amphiphilic ion ch
38	33	84.6	12	1 R36372	Amphiphilic ion ch
39	33	84.6	13	1 R36373	Amphiphilic ion ch
40	33	84.6	14	1 R36374	Amphiphilic ion ch
41	33	84.6	16	1 R36376	Amphiphilic ion ch
42	33	84.6	11	1 R39074	Biologically activ
43	33	84.6	12	1 R39075	Biologically activ

ALIGNMENTS

RESULT 1

R71782  
ID R71782 standard; peptide; 9 AA.  
AC R71782:  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WPI: 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PT amino acids  
PS Claim 13; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 9 AA;

Query Match

100.0%; Score 39; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9

Db 1 KLLKLLKLL 9

RESULT 2

W21599  
ID W21599 standard; peptide; 9 AA.  
AC W21599;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #11.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1995; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M.  
DR WPI: 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 15; Page 25; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to

44 33 84.6 13 1 R39076 Biologically activ  
45 33 84.6 14 1 R39077 Biologically activ

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 39; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||

Db 1 KLLKLLKLL 9

RESULT 3  
 W77378 ID W77378 standard; peptide; 37 AA.

AC W77378;  
 DT 14-DEC-1998 (first entry)  
 DE Lytic peptide with alterable function 3.  
 KW Biologically active peptide; hormone; drug; toxin;  
 KW lipid bilayer membrane; microorganism; parasite; virus.  
 OS Synthetic.  
 PN W09841535-A2.  
 PD 24-SEP-1998.  
 PF 18-MAR-1998; G00799.  
 PR 18-MAR-1997; GB-005519.  
 PA (ANMA-) ANMAT TECHNOLOGY LTD.  
 PI Ajoula HS, Clarke DJ;  
 DR WPI; 98-521161/44.  
 PT New modified peptide(s) - obtained by substitution with an amino  
 PT acid which is modifiable by a reaction and replacing other amino  
 PT acids which are not to be modified.  
 PS Claim 7; Page 22; 33pp; English.  
 CC The peptides W77376-W77390 can be modified by the method of the  
 CC invention by substituting at least one amino acid of the peptide to  
 CC provide a peptide having at least one amino acid which is modifiable by a  
 CC reaction and replacing other amino acids in the peptide with amino acids  
 CC which are not modifiable by the reaction. The methods can be used for  
 CC the modification of biologically active peptides such as hormones, drugs,  
 CC toxins and peptides which act on lipid bilayer membranes. The modified  
 CC peptides can be used e.g. in the body of an animal or plant or parts in  
 CC order to affect the structure or integrity or permeability of a foreign  
 CC body such as a microorganism, parasite or virus present in the body of  
 CC the animal or plant or within the cells of the body of the animal or  
 CC plant.  
 SQ Sequence 37 AA;

Query Match 87.2%; Score 34; DB 1; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 3.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLKLLKLL 9  
 |||||

Db 6 LLKLLKLL 13

RESULT 4  
 P91334 ID P91334 standard; peptide; 28 AA

AC P91334  
 DT 18-MAR-1990 (first entry)  
 DE Amino acid sequence of Shiva-2.  
 KW Shiva-2; lytic peptide; antimicrobial peptide; disease-resistant  
 KW trichophyte; Shiva-3; Shiva-4; Shiva-5; Shiva-6; Shiva-7.  
 PN W08904371-A.  
 PD 18-MAY-1989.  
 PF 2-NOV-1988; U03908.  
 PR 02-NOV-1987; US-115941.

PA (LOUV) Louisiana State Univ.  
 PI Jaynes JM, Derrick KS;  
 DR WPI; 89-165650/22.  
 PT Transformed plants contg. heterologous gene - expressing antimicrobial  
 PS Table I; 56pp; English.  
 CC agent, or polypeptide high in essential amino acids  
 CC Amino acid sequence of Shiva-2 as an exemplary lytic peptide for  
 CC use as an antimicrobial peptide contemplated for use in plant  
 CC (trichophyte) transformants in the invention. It is a homologue of  
 CC Shivas -3 to -7. All of these Shiva peptides are also contemplated as  
 CC having general utility in inducing lysis of cells in vitro.  
 SQ Sequence 28 AA;

Query Match 84.6%; Score 33; DB 1; Length 28;  
 Best Local Similarity 88.9%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||

Db 3 KLLKLLKLL 11

RESULT 5  
 P91335 ID P91335 standard; peptide; 28 AA

AC P91335  
 DT 19-MAR-1990 (first entry)  
 DE Amino acid sequence of Shiva-3.  
 KW Shiva-3; lytic peptide; antimicrobial peptide; disease-resistant  
 KW trichophyte; Shiva-2; Shiva-4; Shiva-5; Shiva-6; Shiva-7.  
 PN W08904371-A.  
 PD 18-MAY-1989.  
 PF 2-NOV-1988; U03908.  
 PR 02-NOV-1987; US-115941.  
 PA (LOUV) Louisiana State Univ.  
 PI Jaynes JM, Derrick KS;  
 DR WPI; 89-165650/22.  
 PT Transformed plants contg. heterologous gene - expressing antimicrobial  
 PS Table I; 56pp; English.  
 CC agent, or polypeptide high in essential amino acids  
 CC Amino acid sequence of Shiva-3 as an exemplary lytic peptide for  
 CC use as an antimicrobial peptide contemplated for use in plant  
 CC (trichophyte) transformants in the invention. It is a homologue of  
 CC Shivas -2, and -4 to -7. All of these Shiva peptides are also  
 CC contemplated as having general utility in inducing lysis of cells in  
 CC vitro. Shiva-3 may be too lytically active to be used in plants at high  
 CC expression levels.  
 SQ Sequence 28 AA;

Query Match 84.6%; Score 33; DB 1; Length 28;  
 Best Local Similarity 88.9%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||

Db 17 KLLKLLKLL 25

RESULT 6

P91337 ID P91337 standard; peptide; 28 AA  
 AC P91337  
 DT 19-MAR-1990 (first entry)  
 DE Amino acid sequence of Shiva-5.  
 KW Shiva-5; lytic peptide; antimicrobial peptide; disease-resistant  
 KW trichophyte; Shiva-2; Shiva-3; Shiva-4; Shiva-6; Shiva-7.  
 PN W08904371-A.  
 PD 18-MAY-1989.  
 PF 2-NOV-1988; U03908.  
 PR 02-NOV-1987; US-115941.  
 PA (LOUV) Louisiana State Univ.



PI Jaynes JM, Derrick KS;  
 DR WPI; 89-165650/22.  
 PT Transformed plants contg. heterologous gene - expressing antimicrobial  
 PS Table I; 56pp; English.  
 CC agent, or polypeptide high in essential amino acids  
 CC Amino acid sequence of Shiva-5 as an exemplary lytic peptide for  
 CC use as an antimicrobial peptide contemplated for use in plant  
 CC (trichophyte) transformants in the invention. It is a homologue of  
 CC Shivas 1-4, 6 and 7. All of these Shiva peptides are also contemplated as  
 CC having general utility in inducing lysis of cells in vitro.  
 SQ Sequence 28 AA;

Query Match 84.6%; Score 33; DB 1; Length 28;  
 Best Local Similarity 88.9%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KLLKLLKLL 9  
 Db 3 KLLKLLKLL 11

RESULT 7  
 ID P91350 standard; peptide; 28 AA  
 AC P91350  
 DT 19-MAR-1990 (first entry)  
 DE Amino acid sequence of Degrado Act.  
 KW Degrado Act.; lytic peptide; antimicrobial peptide; disease-resistant  
 KW trichophyte.  
 PN W08904371-A.  
 PD 18-MAY-1989.  
 PF 2-NOV-1988; U03908.  
 PR 02-NOV-1987; US-115941.  
 PA (LOU) Louisiana State Univ.  
 PI Jaynes JM, Derrick KS;  
 DR WPI; 89-165650/22.  
 PT Transformed plants contg. heterologous gene - expressing antimicrobial  
 PT agent, or polypeptide high in essential amino acids  
 PS Table I; 56pp; English.  
 CC Amino acid sequence of Degrado Act. as an exemplary lytic peptide  
 CC for use as an antimicrobial peptide contemplated for use in plant  
 CC (trichophyte) transformants in the invention.  
 SQ Sequence 28 AA;

Query Match 84.6%; Score 33; DB 1; Length 28;  
 Best Local Similarity 88.9%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KLLKLLKLL 9  
 Db 3 KLLKLLKLL 11

RESULT 8  
 ID R21386 standard; peptide; 17 AA.  
 AC R21386;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 49 with  
 DE C-terminal amide and acetylated at N-terminus.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; sterilant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating

PT viral and phytopathogenic infections, tumours and burns  
 PS Claim 4; Page 54; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing.  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 17 AA;

Query Match 84.6%; Score 33; DB 1; Length 17;  
 Best Local Similarity 88.9%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 Db 2 KLLKLLKLL 10

RESULT 9  
 ID R21388 standard; peptide; 17 AA.  
 AC R21388;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 51 with  
 DE C-terminal amide and acetylated at N-terminus.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; sterilant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PR 17-JUL-1991; U05047.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating  
 PT viral and phytopathogenic infections, tumours and burns  
 PS Claim 4; Page 55; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing.  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 17 AA;

Query Match 84.6%; Score 33; DB 1; Length 17;  
 Best Local Similarity 88.9%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KLLKLLKLL 9  
 Db 6 KLLKLLKLL 14

RESULT 10  
 ID R21390 standard; peptide; 17 AA.  
 AC R21390;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 53 with  
 DE C-terminal amide and acetylated at N-terminus.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; sterilant.

PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating  
 PT viral and phytopathogenic infections, tumours and burns  
 PS Claim 4; Page 56; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing,  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 17 AA:

Query Match 84.6%; Score 33; DB 1; Length 17;  
 Best Local Similarity 88.9%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||  
 Db 6 KLLKLLKLL 14

RESULT 11  
 R21397  
 ID R21397 standard; peptide; 22 AA.  
 AC R21397;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 60 with  
 DE acetylated N-terminus, may be a C-terminal amide.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; steriliant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating  
 PT viral and phytopathogenic infections, tumours and burns  
 PS Disclosure; Page 59; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing,  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 22 AA:

Query Match 84.6%; Score 33; DB 1; Length 22;  
 Best Local Similarity 88.9%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||  
 Db 3 KLLKLLKLL 11

RESULT 12  
 R21398  
 ID R21398 standard; peptide; 26 AA.  
 AC R21398;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 61 with  
 DE acetylated N-terminus, may be a C-terminal amide.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; steriliant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating  
 PT viral and phytopathogenic infections, tumours and burns  
 PS Disclosure; Page 60; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing,  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 26 AA:

Query Match 84.6%; Score 33; DB 1; Length 26;  
 Best Local Similarity 88.9%; Pred. No. 3.8;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLLKLLKLL 9  
 |||||  
 Db 7 KLLKLLKLL 15

RESULT 13  
 R21399  
 ID R21399 standard; peptide; 30 AA.  
 AC R21399;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID No. 62 with  
 DE acetylated N-terminus, may be a C-terminal amide.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour;  
 KW spermicide; wound healing; steriliant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises  
 PT administering amphiphilic peptide compns, useful for treating  
 PT viral and phytopathogenic infections, tumours and burns  
 PS Disclosure; Page 60; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having  
 CC anti-microbial, anti-viral and anti-tumour activity. They are also  
 CC useful for inhibiting, preventing or destroying the motility of  
 CC sperm and hence have application in a spermicide preparation. They  
 CC also have anti-parasitic activity and are useful in wound healing,  
 CC as preservatives and sterilants and to inhibit growth of  
 CC phytopathogenic microorganisms. R20969 and R20970 were published in  
 CC Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83,  
 CC 1987.  
 SQ Sequence 30 AA:

Query Match 84.6%; Score 33; DB 1; Length 30;  
 Best Local Similarity 88.9%; Pred. No. 4.4; Mismatches 0; Indels 1; Gaps 0;  
 Matches 8; Conservative 0

QY 1 KLLKLLKLL 9  
 ||||| |  
 DB 11 KLLKLLKLL 19

## RESULT 14

R21400 ID R21400 standard; peptide; 36 AA.  
 AC R21400;  
 DT 16-MAY-1992 (first entry)  
 DE Sequence of amphiphilic peptide SEQ ID NO. 63 with acetylated N-terminus, may be a C-terminal amide.  
 KW Amphiphilic peptide; anti-microbial; anti-viral; anti-tumour; spermicide; wound healing; sterilant.  
 PN W09201462-A.  
 PD 06-FEB-1992.  
 PF 17-JUL-1991; U05047.  
 PR 19-JUL-1990; US-554422.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-064700/08.  
 PT Method for inhibiting target cell and virus growth - comprises administering amphiphilic peptide compns, useful for treating viral and phytopathogenic infections, tumours and burns  
 PS Disclosure; Page 61; 72pp; English.  
 CC The peptides of the invention are effective pharmaceuticals having anti-microbial, anti-viral and anti-tumour activity. They are also useful for inhibiting, preventing or destroying the motility of sperm and hence have application in a spermicide preparation. They also have anti-parasitic activity and are useful in wound healing, as preservatives and sterilants and to inhibit growth of phytopathogenic microorganisms. R20969 and R20970 were published in Houghten and Ostresh, Bio Chromatography, Vol 2, issue 2, page 80-83, 1987.  
 CC 1987.  
 SQ Sequence 36 AA;

Query Match 84.6%; Score 33; DB 1; Length 36;  
 Best Local Similarity 88.9%; Pred. No. 5.2; Mismatches 0; Indels 1; Gaps 0;  
 Matches 8; Conservative 0

QY 1 KLLKLLKLL 9  
 ||||| |  
 DB 17 KLLKLLKLL 25

## RESULT 15

R22870 ID R22870 standard; Peptide; 17 AA.  
 AC R22870;  
 DT 22-AUG-1992 (first entry)  
 DE Amphiphilic peptide to inhibit growth of a target cell.  
 KW Antimicrobial; antifungal; antitumour; pharmaceuticals; antibiotics; antiparasitic; spermicides; burns; wound healing.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT modified\_site 1 /note= "acylated"  
 FT modified\_site 17 /note= "amidated"  
 FT modified\_site 17 /note= "amidated"  
 PN CA2047317-A.  
 PD 20-JAN-1992.  
 PF 18-JUL-1991; 024317.  
 PR 19-JUL-1990; US-554442.  
 PR 08-JUL-1991; US-725331.  
 PA (SCRI-) SCRIPPS RES INST.  
 PI Houghten RA, Blondelle S;  
 DR WPI; 92-114943/15.  
 PT Amphiphilic peptide(s) and analogues - for use in e.g.

PT antimicrobial, antifungal or antitumour compositions, having increased biological activity  
 PS Claim 9; Page 54; 71pp; English.  
 CC The amphiphilic peptide (SEQ ID NO 49) was prepd. by standard solid phase methods. The peptide is an analogue of the peptide having SEQ ID NO 3 (R22824), with Lys at position 5 deleted.  
 CC Substitution and deletion analogues of this peptide have increased biological activity and are effective as pharmaceuticals e.g. antibiotics for bacterial, fungal or viral infections, or in spermicides or antitumour or antiparasitic agents. Additionally the peptides can be used in wound healing compns. or for treating burns or other skin or eye infections.  
 CC See also R22822-89.  
 SQ Sequence 17 AA;

Query Match 84.6%; Score 33; DB 1; Length 17;  
 Best Local Similarity 88.9%; Pred. No. 2.5; Mismatches 0; Conservative 0; Indels 1; Gaps 0;  
 Matches 8; Conservative 0

QY 1 KLLKLLKLL 9  
 ||||| |  
 DB 2 KLLKLLKLL 10

Search completed: September 7, 1999, 20:37:13  
 Job time: 18462 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: September 7, 1999, 20:34:42 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-12  
Perfect score: 52  
Sequence: 1 KKKKKKFKK 10  
Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL\_10:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	84.6	287	13	O93503 xenopus lae
2	41	78.8	369	5	Q21024 caenorhabdi
3	39	75.0	191	5	O45682 caenorhabdi
4	39	75.0	1174	5	Q26240 rhodnius pr
5	38	73.1	249	2	Q92H28 mycoplasma
6	38	73.1	1189	4	O43604 homo sapien
7	38	73.1	1199	4	O60668 homo sapien
8	38	73.1	1199	4	O43487 homo sapien
9	38	73.1	1247	5	O96168 plasmodium
10	38	73.1	1542	5	O97299 plasmodium
11	37	71.2	539	4	O75322 homo sapien
12	37	71.2	261	5	O18662 plasmodium
13	37	71.2	444	10	P93716 petunia hyb
14	37	71.2	129	11	O35807 ratius norv
15	36	69.2	295	1	O58667 pyrococcus
16	36	69.2	659	1	O57721 pyrococcus
17	36	69.2	1116	2	O67838 aquifex aeo
18	36	69.2	114	2	O57459 coxiella bu
19	36	69.2	815	4	O43273 homo sapien
20	36	69.2	655	4	O15311 homo sapien
21	36	69.2	473	4	O75949 homo sapien
22	36	69.2	170	5	O44172 caenorhabdi
23	36	69.2	1421	5	O19625 caenorhabdi
24	36	69.2	805	5	O45742 caenorhabdi
25	36	69.2	2485	5	O96134 plasmodium
26	36	69.2	1516	5	O96154 plasmodium
27	36	69.2	2013	5	O96216 plasmodium
28	36	69.2	1817	5	O96253 plasmodium
29	36	69.2	1569	5	O97234 plasmodium

30	36	69.2	1096	5	O97257 plasmodium
31	36	69.2	756	10	O65716 arabidopsis
32	36	69.2	204	10	O23504 arabidopsis
33	36	69.2	431	10	O80587 arabidopsis
34	36	69.2	796	10	O49464 arabidopsis
35	36	69.2	693	11	O55195 rattus norv
36	36	69.2	648	11	O62172 mus musculu
37	36	69.2	647	11	O62796 rattus norv
38	36	69.2	213	11	P97762 mus musculu
39	36	69.2	188	13	O42277 oryzias lat
40	36	69.2	188	13	O93280 pleuronecte
41	36	69.2	188	13	O9VH38 cyprinus ca
42	35	67.3	629	2	O06476 bacillus su
43	35	67.3	1051	3	Q01694 pneumocysti
44	35	67.3	999	3	Q09038 saccharomyc
45	35	67.3	102	11	O63672 rattus norv

ALIGNMENTS

RESULT 1  
O93503 PRELIMINARY; PRT; 287 AA.  
AC O93503;  
DT 01-NOV-1998 (TREMELREL. 08, Created)  
DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)  
DT 01-MAY-1999 (TREMELREL. 10, Last annotation update)  
DE MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE.  
GN MARCKS.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;  
OC Xenopus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98030614.  
RA SHI Y., SULLIVAN S.K., PITTLERLE D.M., KENNINGTON E.A., GRAFF J.M.,  
RA BLACKSHEAR P.J.;  
RL J. Biol. Chem. 272:29290-29300(1997).  
RL EMBL; AF017299; AAC61897.1; .  
DR PROSITE; PS00826; MARCKS\_1; 1.  
DR PROSITE; PS00827; MARCKS\_2; 1.  
SQ SEQUENCE 287 AA; 29147 MW; 506899EEE CRC32;

Query Match 84.6%; Score 44; DB 13; Length 287;  
Best Local Similarity 80.0%; Pred. No. 10;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKKKKKFKK 10  
|||||  
DB 125 KKKKKRFSK 134

RESULT 2

Q21024 PRELIMINARY; PRT; 369 AA.  
ID Q21024;  
DT 01-NOV-1996 (TREMELREL. 01, Created)  
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)  
DT 01-NOV-1998 (TREMELREL. 08, Last annotation update)  
DE SIMILAR TO RIBONUCLEASE H.  
GN P59A5.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,

```

RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA NHAN M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41994; AAA83453.1; -.
DR PFAM; PF00075; fnaseH; 1.
SQ SEQUENCE 369 AA; 42109 MW; B6C92623 CRC32;

Query Match 78.8%; Score 41; DB 5; Length 369;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFK 8
Db 149 KKKKKKFK 156
|||||||

RESULT 3
ID O45682 PRELIMINARY; PRT; 191 AA.
AC O45682;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE K10H10.5 PROTEIN.
GN K10H10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA PERCY C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 283236; CAB05781.1; -.
SQ SEQUENCE 191 AA; 21158 MW; ADEA462A CRC32;

Query Match 75.0%; Score 39; DB 5; Length 191;
Best Local Similarity 77.8%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 283236; CAB05781.1; -.
SQ SEQUENCE 191 AA; 21158 MW; ADEA462A CRC32;

Query Match 75.0%; Score 39; DB 5; Length 1174;
Best Local Similarity 70.0%; Pred. No. 1.6e-02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFK 10
Db 473 KKKKKKFK 482
|||||||

RESULT 5
ID Q92H28 PRELIMINARY; PRT; 249 AA.
AC Q92H28;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE 30 KDA GTP-BINDING PROTEIN LEPA.
OS Mycoplasma hyopneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

```

```

QY 1 KKKKKKFK 9
Db 71 KKKKKKFR 79
|||||||

RESULT 4
ID Q26240 PRELIMINARY; PRT; 1174 AA.
AC Q26240;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE NITRIC-OXIDE SYNTHASE (EC 1.14.13.39).
OS Rhodnius prolixus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Hemiptera; Euhemiptera; Heteroptera; Panheteroptera;
OC Cimicomorpha; Reduviidae; Triatominae; Rhodnius.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLAND;
RX MEDLINE; 97175053.
RA YUDA M., HIRAI M., MIURA K., MATSUMURA H., ANDO K., CHINZEI Y.;
RT "cDNA cloning, expression and characterization of nitric-oxide
RT synthase from the salivary glands of the blood-sucking insect
RT Rhodnius prolixus.";
RL Eur. J. Biochem. 242:807-812(1996).
CC -!- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE
CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY. THE PRODUCTION OF NO
CC IN THE SALIVARY GLAND IS USED AS A VASODILATOR FOR BLOOD SUCKING.
CC -!- CATALYTIC ACTIVITY: L-ARGININE + N NADPH + M O(2) = CITRULLINE +
CC NITRIC OXIDE + N NADP(+).
CC -!- COFACTOR: BINDS ONE MOLE EACH OF FAD AND FMN AND ALSO BINDS
CC TETRAHYDROBIOTERIN (BY SIMILARITY).
CC -!- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: STRONG, TO OTHER NOS ISOZYMES. ALSO TO CYTOCHROME
CC P-450 REDUCTASE.
DR EMBL; U59389; AAB03810.1; -.
DR PFAM; PF00175; oxidorel_fad; 1.
DR PFAM; PF00667; FAD_binding; 1.
KW Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Calcium-binding;
KW Heme.
FT DOMAIN 35 45 POLY-GLN.
FT BINDING 162 162 HEME (BY SIMILARITY).
FT DOMAIN 316 357 ARGININE-BINDING (BY SIMILARITY).
FT DOMAIN 475 495 CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 639 670 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 780 791 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 923 933 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 998 1016 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1095 1110 NADP (ADP PART) (BY SIMILARITY).
SQ SEQUENCE 1174 AA; 132392 MW; CC523CE8 CRC32;

Query Match 75.0%; Score 39; DB 5; Length 1174;
Best Local Similarity 70.0%; Pred. No. 1.6e-02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFK 10
Db 473 KKKKKKFK 482
|||||||

RESULT 5
ID Q92H28 PRELIMINARY; PRT; 249 AA.
AC Q92H28;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE 30 KDA GTP-BINDING PROTEIN LEPA.
OS Mycoplasma hyopneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

```

OC Mycoplasmataceae; Mycoplasma.

RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=232;  
RA LIN H.N., SHIUAN D.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF046228; AAC98966.1; -;  
DR PROSITE; PS00301; EFACOR\_GTP; 1.  
KW Protein biosynthesis; GTP-binding.  
SQ SEQUENCE 249 AA; 28691 MW; 9250CB2F CRC32;

Query Match 73.1%; Score 38; DB 2; Length 249;  
Best Local Similarity 77.8%; Pred. No. 65;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKKKFKK 10  
||| |||:  
Db 6 KKKPKFKYK 14

RESULT 6  
O43604 PRELIMINARY; PRT; 1189 AA.  
AC O43604;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE TATA BINDING PROTEIN ASSOCIATED FACTOR.  
GN TAFI1150.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MARTINEZ E., GE H., TAO Y., YUAN C.-X., ROEDER R.G.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF040701; AAC68502.1; -;  
SQ SEQUENCE 1189 AA; 135833 MW; 180759F2 CRC32;

Query Match 73.1%; Score 38; DB 4; Length 1189;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
||||| |  
Db 1145 KKKKKKHK 1154

RESULT 7  
O60668 PRELIMINARY; PRT; 1199 AA.  
AC O60668;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE TBP-ASSOCIATED FACTOR TAFI1150.  
GN TAFI1150.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GUERMAH M., ROEDER R.G.R.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF057694; AAC13540.1; -;  
SQ SEQUENCE 1199 AA; 136985 MW; 98118DA8 CRC32;

Query Match 73.1%; Score 38; DB 4; Length 1199;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
||||| |  
Db 1155 KKKKKKHK 1164

RESULT 8  
O43487 PRELIMINARY; PRT; 1199 AA.  
AC O43487;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE COFACTOR OF INITIATOR FUNCTION.  
GN CIFI50.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98078679.  
RA KAUFMANN J., AHRENS K., KOOP R., SMALE S.T., MULLER R.;  
RT "CIFI50, a human cofactor for transcription factor IID-dependent  
RT initiator function."  
RL Mol. Cell. Biol. 18:233-239(1998).  
DR EMBL: AF026445; AAC02966.1; -;  
SQ SEQUENCE 1199 AA; 136993 MW; A6363760 CRC32;

Query Match 73.1%; Score 38; DB 4; Length 1199;  
Best Local Similarity 80.0%; Pred. No. 2.2e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
||||| |  
Db 1155 KKKKKKHK 1164

RESULT 9  
O96168 PRELIMINARY; PRT; 1247 AA.  
AC O96168;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 149.3 KD PROTEIN.  
GN PFB0365W.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
RT falciparum."  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001390; AAC71862.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 1247 AA; 149260 MW; B8B41F66 CRC32;

Query Match 73.1%; Score 38; DB 5; Length 1247;  
Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 8  
||:||||  
Db 483 KKKKKKFK 490

```

RESULT 10
O97299 ID O97299 PRELIMINARY; PRT; 1542 AA.
AC O97299;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE MAL3P7.37 PROTEIN.
GN MAL3P7.37.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AL034559; CAB39046.1; -.
SQ SEQUENCE 1542 AA; 184460 MW; 242A4DE9 CRC32;

Query Match 73.1%; Score 38; DB 5; Length 1542;
Best Local Similarity 77.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9
DB 26 KKKKKLKY 34
||||| |

RESULT 11
O75322 ID O75322 PRELIMINARY; PRT; 539 AA.
AC O75322;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE HSP89-ALPHA-DELTA-N.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98267138.
RA SCHWEINFEST C.W., GRABER M.W., HENDERSON K.W., PAPAS T.S., BARON P.L.,
RA WATSON D.K.;
RT "Cloning and sequence analysis of Hsp89alpha Deltan, a new member of
RT theHsp90 gene family.";
RL Biochim. Biophys. Acta 1398:18-24(1998).
DR ENBL; AF028832; AAC25497.1; -.
DR PFAM; PF00183; HSP90; 1.
SQ SEQUENCE 539 AA; 63251 MW; 27092319 CRC32;

Query Match 71.2%; Score 37; DB 4; Length 539;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
DB 81 KKKKKKIKK 90
||||| |

RESULT 12
O18662 ID O18662 PRELIMINARY; PRT; 261 AA.
AC O18662;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE CG9 PROTEIN.
GN CG9.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD2, AND HB3;
RX MEDLINE; 98054002.
RA SU X.-Z., KIRKMAN L.A., FUJIOKA H., WELLEMS T.E.;
RT "Complex polymorphisms in an approximately kDa protein are linked to
RT chloroquine-resistant P. falciparum in Southeast Asia and Africa.";
RL Cell 91:593-603(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DD2, AND HB3;
RA SU X.-Z., KIRKMAN L.A., WELLEMS T.E.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF030694; AAC47839.1; -.
DR ENBL; AF030690; AAC47852.1; -.
SQ SEQUENCE 261 AA; 31108 MW; 5CFBDC37 CRC32;

Query Match 71.2%; Score 37; DB 5; Length 261;
Best Local Similarity 80.0%; Pred. No. 93;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
DB 233 KKKKKKKIK 242
||||| |

RESULT 13
P93716 ID P93716 PRELIMINARY; PRT; 444 AA.
AC P93716;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE PETHY, ZPT3-2.
OS Petunia hybrida (Petunia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;
OC Petunia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MITCHELL DIPLOID; TISSUE=STAMEN;
RA KOBAYASHI A., SAKAMOTO A., KUBO K., RYBKA Z., KANNO Y., TAKATSUJI H.;
RT "Seven zinc-finger transcription factors are expressed sequentially
RT during the development of anthers and pollen in Petunia.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB000454; BAA19113.1; -.
DR MENDEL; L2065; PEX; 1523; 4.
DR PFAM; PF00096; zf-C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
KW zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 444 AA; 49268 MW; E7D50BB0 CRC32;

Query Match 71.2%; Score 37; DB 10; Length 444;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
DB 264 KKKKKIKLR 273
||||| |

RESULT 14
O35807 ID O35807 PRELIMINARY; PRT; 129 AA.
AC O35807;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE MICROVASCULAR ENDOTHELIAL DIFFERENTIATION GENE 2.
GN MDG2.

```



OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EPIDIDYMAL;  
 RC PROELS F., LOSER B., MARX M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y08769; CAA70022.1; -;  
 DR PFAM; PF00069; pKinase; 1.  
 SQ SEQUENCE 129 AA; 15080 MW; 7A874092 CRC32;

Query Match 71.2%; Score 37; DB 11; Length 129;  
 Best Local Similarity 80.0%; Pred. No. 54;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KKKKKKFK 10  
 DB 98 KKKKKKKIK 107  
 ||||| |

RESULT 15  
 C58657 PRELIMINARY: PRT: 295 AA.  
 ID 058667  
 AC 058667;  
 DT 01-AUG-1998 (TReMBLrel. 07, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TReMBLrel. 09, Last annotation update)  
 DE 295AA LONG HYPOTHETICAL MRP PROTEIN.  
 GN PH0949.  
 OS Pyrococcus horikoshii.  
 RC Archaea; Euryarchaeota; Thermococcales; Pyrococcaceae; Pyrococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE: 98344137.  
 RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSUYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
 RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT "Complete sequence and gene organization of the genome of a  
 RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL; AP000004; BAA30046.1; -;  
 SQ SEQUENCE 295 AA; 32033 MW; 531A7C76 CRC32;

Query Match 69.2%; Score 36; DB 1; Length 295;  
 Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KKKKKKFK 10  
 DB 23 KEKERRKKYK 32  
 |:|:|:|

Search completed: September 7, 1999, 20:34:43  
 Job time: 19739 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:15 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-12

Perfect score: 52

Sequence: 1 KKKKKKFKK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	84.6	331	1 MACS_BOVIN	P12624 bos taurus
2	44	84.6	280	1 MACS_CHICK	P16527 gallus gall
3	44	84.6	331	1 MACS_HUMAN	P29966 homo sapien
4	44	84.6	308	1 MACS_MOUSE	P26645 mus musculus
5	44	84.6	308	1 MACS_RAT	P30009 rattus norv
6	42	80.8	194	1 MRP_HUMAN	P49006 homo sapien
7	42	80.8	199	1 MRP_MOUSE	P28667 mus musculus
8	42	80.8	198	1 MRP_RABIT	P35566 oryctolagus
9	39	75.0	114	1 VG40_BPT4	P17171 bacteriophag
10	38	73.1	830	1 PAM1_YEAST	P37304 saccharomyc
11	37	71.2	728	1 HS9A_CHICK	P11501 gallus gall
12	37	71.2	732	1 HS9A_CRIGR	P46633 cricetus
13	37	71.2	721	1 HS9A_HUMAN	P07900 homo sapien
14	37	71.2	732	1 HS9A_MOUSE	P07901 mus musculus
15	37	71.2	732	1 HS9A_PIG	O02705 sus scrofa
16	36	69.2	188	1 RASK_MELGA	P79800 meleagris g
17	36	69.2	188	1 RASK_MONDO	O07983 monodelphis
18	36	69.2	188	1 RASK_HUMAN	P01118 homo sapien
19	36	69.2	312	1 YG37_YEAST	P53275 saccharomyc
20	35	67.3	370	1 CTPT_PLAFK	P49587 plasmodium
21	35	67.3	514	1 DKC1_HUMAN	O60832 homo sapien
22	35	67.3	333	1 IF2B_HUMAN	P20042 homo sapien
23	35	67.3	333	1 IF2B_RABIT	P41035 oryctolagus
24	35	67.3	209	1 NUGM_PAPPR	P15601 parametium
25	35	67.3	209	1 NUGM_PAPTE	P15600 parametium
26	35	67.3	396	1 REPA_BACSU	P13962 bacillus su
27	35	67.3	141	1 RS12_METH	O27129 methanobact
28	35	67.3	214	1 VEGE_MOUSE	O00731 mus musculus
29	35	67.3	3951	1 VGFL_IBVB	P27920 avian infec
30	35	67.3	224	1 Y364_MYCGE	P47604 mycoplasma
31	35	67.3	2136	1 YCF2_MARPO	P09975 marchantia
32	35	67.3	295	1 YCX7_EUGGR	P31920 euglena gra
33	34	65.4	686	1 CNG1_HUMAN	P29973 homo sapien
34	34	65.4	291	1 DTCM_MOUSE	P01882 mus musculus
35	34	65.4	258	1 DTC_MOUSE	P01881 mus musculus
36	34	65.4	508	1 NO60_DROME	O44081 drosophila
37	34	65.4	441	1 PHPA_PLACH	O02752 plasmodium
38	34	65.4	418	1 RMS1_HUMAN	P49646 homo sapien
39	34	65.4	1562	1 RPOD_CHLUV	P12455 chlorella v
40	34	65.4	2233	1 RRPL_PT3H4	P12677 human paral
41	34	65.4	670	1 SR72_CANFA	P33731 canis famil
42	34	65.4	670	1 SR72_HUMAN	O76094 homo sapien
43	34	65.4	1213	1 T2D2_DROME	O24325 drosophila

44 34 65.4 322 1 VG3\_SPVIR P15894 spiroplasma  
45 34 65.4 160 1 YCX2\_ASTLO P34777 astasia lon

## ALIGNMENTS

RESULT 1  
MACS\_BOVIN STANDARD; PRT; 331 AA.  
ID MACS\_BOVIN  
AC P12624;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS) (ACAMP-81).  
GN MACS.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89282412.  
RA STUMPO D.J., GRAFF J.M., ALBERT K.A., GRENGARD P., BLACKSHEAR P.J.;  
RT "Nucleotide sequence of a cDNA for the bovine myristoylated  
alanine-rich C kinase substrate (MARCKS).";  
RL NUCLEIC ACIDS RES. 17:3987-3988(1989).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 89264553.  
RA STUMPO D.J., GRAFF J.M., ALBERT K.A., GRENGARD P., BLACKSHEAR P.J.;  
RT "Molecular cloning, characterization, and expression of a cDNA  
encoding the '80- to 87-kDa' myristoylated alanine-rich C kinase  
RT substrate: a major cellular substrate for protein kinase C.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:4012-4016(1989).  
RN [3]  
RP PARTIAL SEQUENCE.  
RX MEDLINE; 92171958.  
RA MIZUTANI A., TOKUMITSU H., HIDAKA H.;  
RT "Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein  
interacting with synapsin I.";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 182:1395-1401(1992).  
RN [4]  
RP PHOSPHORYLATION SITES.  
RX MEDLINE; 89308594.  
RA GRAFF J.M., STUMPO D.J., BLACKSHEAR P.J.;  
RT "Characterization of the phosphorylation sites in the chicken and  
bovine myristoylated alanine-rich C kinase substrate protein, a  
RT prominent cellular substrate for protein kinase C.";  
RL J. BIOL. CHEM. 264:11912-11919(1989).  
RN [5]  
RP PHOSPHORYLATION SITES, AND REVISIONS.  
RX TISSUE=BRAIN;  
MEDLINE; 94308052.  
RA TANIGUCHI H., MANENTI S., SUZUKI M., TITANI K.;  
RT "Myristoylated alanine-rich C kinase substrate (MARCKS), a major  
RT protein kinase C substrate, is an in vivo substrate of  
RT proline-directed protein kinase(s). A mass spectroscopic analysis of  
RT the post-translational modifications.";  
RL J. BIOL. CHEM. 269:18299-18302(1994).  
RN [6]  
RP REVERSIBLE ASSOCIATION WITH THE MEMBRANE.  
RX MEDLINE; 91238951.  
RA THELEN M., ROSEN A., NAIRN A.C., ADEREM A.;  
RT "Regulation by phosphorylation of reversible association of a  
RT myristoylated protein kinase C substrate with the plasma membrane.";  
RL NATURE 351:320-322(1991).  
RN [7]  
RP ACTIN-FILAMENT CROSS-LINKING.  
RX MEDLINE; 92220195.  
RA HARTWIG J.H., THELEN M., ROSEN A., JANNEY P.A., NAIRN A.C.,  
RA ADEREM A.;  
RT "MARCKS is an actin filament crosslinking protein regulated by  
RT protein kinase C and calcium-calmodulin.";

```

RL NATURE 356:618-622(1992).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC -!- CAUTION: REF.1 AND REF.2 SEQUENCES DIFFER FROM THAT SHOWN IN
CC POSITIONS 140 TO 150 DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M24638; G163340; ALT_FRAME.
CC PIR: A32904; A32904.
CC PIR: S03338; S03338.
CC PIR: S08341; S08341.
CC PROSITE: PS00826; MARCKS_1; 1.
CC PROSITE: PS00827; MARCKS_2; 1.
CC PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;
CC MEMBRANE.
CC INIT_MET 0 0
CC LIPID 1 1 MYRISTATE.
CC DOMAIN 150 174 CALMODULIN-BINDING (PSD).
CC MOD_RES 26 26 PHOSPHORYLATION.
CC MOD_RES 45 45 PHOSPHORYLATION.
CC MOD_RES 80 80 PHOSPHORYLATION.
CC MOD_RES 99 99 PHOSPHORYLATION.
CC MOD_RES 116 116 PHOSPHORYLATION.
CC MOD_RES 133 133 PHOSPHORYLATION.
CC MOD_RES 157 157 PHOSPHORYLATION (BY PKC).
CC MOD_RES 161 161 PHOSPHORYLATION (BY PKC).
CC MOD_RES 165 165 PHOSPHORYLATION (BY PKC).
CC MOD_RES 168 168 PHOSPHORYLATION (BY PKC).
CC SEQUENCE 331 AA; 31450 MW; 27695959 CRC32;

Query Match 84.6%; Score 44; DB 1; Length 331;
Best Local Similarity 80.0%; Pred. No. 1.8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
| | | | | | | |
DB 150 KKKKKKFSFK 159

RESULT 2
MACS_CHICK STANDARD; PRT; 280 AA.
AC P16527;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE; 90114197.
RA GRAFF J.M., STUMPO D.J., BLACKSHEAR P.J.;
RT "Molecular cloning, sequence, and expression of a cDNA encoding the
RT chicken myristoylated alanine-rich C kinase substrate (MARCKS).";
RL MOL. ENDOCRINOL. 3:1903-1906(1989).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT

```

```

CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M31650; G212288; -.
CC PIR: A41400; A41400.
CC PROSITE: PS00826; MARCKS_1; 1.
CC PROSITE: PS00827; MARCKS_2; 1.
CC HSP: P04002; LWFA.
CC PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;
CC MEMBRANE.
CC INIT_MET 0 0
CC LIPID 1 1 MYRISTATE.
CC DOMAIN 116 140 CALMODULIN-BINDING (PSD).
CC MOD_RES 123 123 PHOSPHORYLATION (BY PKC).
CC MOD_RES 127 127 PHOSPHORYLATION (BY PKC).
CC MOD_RES 131 131 PHOSPHORYLATION (BY PKC).
CC MOD_RES 134 134 PHOSPHORYLATION (BY PKC).
CC SEQUENCE 280 AA; 27597 MW; 41D34538 CRC32;

Query Match 84.6%; Score 44; DB 1; Length 280;
Best Local Similarity 80.0%; Pred. No. 1.5;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
| | | | | | | |
DB 116 KKKKKKFSFK 125

RESULT 3
MACS_HUMAN STANDARD; PRT; 331 AA.
AC P29966;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS) (PROTEIN KINASE
DE C SUBSTRATE, 80 KD PROTEIN, LIGHT CHAIN) (PKCSL) (80K-L PROTEIN).
GN MACS.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91317795.
RA HARLAN D.M., GRAFF J.M., STUMPO D.J., EDDY R.L. JR., SHOWS T.B.,
RA BOYLE J.M., BLACKSHEAR P.J.;
RT "The human myristoylated alanine-rich C kinase substrate (MARCKS)
RT gene (MACS). Analysis of its gene product, promoter, and chromosomal
RT localization.";
RL J. BIOL. CHEM. 266:14399-14405(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93052291.
RA SAKAI K., HIRAI M., KUDOH J., MINOSHIMA S., SHIMIZU N.;
RT "Molecular cloning and chromosomal mapping of a cDNA encoding human
RT 80K-L protein: major substrate for protein kinase C.";
RL GENOMICS 14:175-178(1992).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M68956; G187387; -;  
 DR EMBL: M68955; G187385; -;  
 DR EMBL: D10532; G219894; -;  
 DR PIR: A38873; A38873;  
 DR MIM: I77061; -;  
 DR PROSITE: PS00826; MARCKS\_1; 1;  
 DR PROSITE: PS00827; MARCKS\_2; 1;  
 KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
 KW MEMBRANE.

FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT DOMAIN 151 175 CALMODULIN-BINDING (PSD).  
 FT MOD\_RES 158 158 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 166 166 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 169 169 PHOSPHORYLATION (BY PKC).  
 FT CONFLICT 83 83 S -> A (IN REF. 1).  
 FT CONFLICT 118 118 A -> P (IN REF. 1).  
 FT CONFLICT 233 233 P -> S (IN REF. 1).  
 FT CONFLICT 286 307 LVCPRGGSPRGARGRRSLNQ (IN REF. 1).  
 FT SEQUENCE 331 AA; 31413 MW; 152CCC42 CRC32;

Query Match 84.6%; Score 44; DB 1; Length 331;  
 Best Local Similarity 80.0%; Pred. No. 1.8;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKKKKFKFK 10  
 |||||:|  
 DB 151 KKKKKRFSEK 160

RESULT 4  
 MACS\_MOUSE STANDARD; PRT; 308 AA.  
 AC P26645;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)  
 DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).  
 GN MACS.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=MACROPHAGE;  
 RX MEDLINE: 91172836.  
 RA SEYKORA J.T., RAVETICH J.V., ADEREM A.;  
 RT "Cloning and molecular characterization of the murine macrophage '68-  
 RT kda' protein kinase C substrate and its regulation by bacterial  
 RT lipopolysaccharide."  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:2505-2509(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=FIBROBLAST;  
 RC MEDLINE: 91330872.  
 RX BROOKS S.F., HERGET T., ERUSALIMSKY J.D., ROZENGURT E.;  
 RT "Protein kinase C activation potentially down-regulates the expression  
 RT of its major substrate, 80K, in Swiss 3T3 cells."  
 RL EMBO J. 10:2497-2505(1991).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RP STRAIN=SWISS; TISSUE=FIBROBLAST;

RX MEDLINE: 90345162.  
 RA BROOKS S.F., ERUSALIMSKY J.D., TOTTY N.F., ROZENGURT E.;  
 RT "Purification and internal amino acid sequence of the 80 kDa protein  
 RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with  
 RT substrates from brain."  
 RL FEBS LETT. 268:291-295(1990).  
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
 CC -!- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND  
 CC VERY LOW LEVELS IN LIVER.  
 CC -!- INDUCTION: BY LIPOPOLYSACCHARIDE.  
 CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
 CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M60474; G199027; -;  
 DR PIR: A39169; A39169.  
 DR PIR: S16519; S16519.  
 DR MGD: MGI:96907; MACS.  
 DR PROSITE: PS00826; MARCKS\_1; 1;  
 DR PROSITE: PS00827; MARCKS\_2; 1;  
 KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
 KW MEMBRANE.

FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT DOMAIN 144 168 CALMODULIN-BINDING (PSD).  
 FT MOD\_RES 151 151 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 155 155 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 159 159 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT CONFLICT 95 97 AGA -> TGT (IN REF. 2).  
 FT SEQUENCE 308 AA; 29530 MW; A91BD349 CRC32;

Query Match 84.6%; Score 44; DB 1; Length 308;  
 Best Local Similarity 80.0%; Pred. No. 1.7;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKKKKFKFK 10  
 |||||:|  
 DB 144 KKKKKRFSEK 153

RESULT 5  
 MACS\_RAT STANDARD; PRT; 308 AA.  
 AC P30009;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
 DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).  
 GN MACS.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE: 91201362.  
 RA ERUSALIMSKY J.D., BROOKS S.F., HERGET T., MORRIS C., ROZENGURT E.;  
 RT "Molecular cloning and characterization of the acidic 80-kDa protein  
 RT kinase C substrate from rat brain. Identification as a  
 RT glycoprotein."  
 RL J. BIOL. CHEM. 266:7073-7080(1991).

RN [2]  
RP PHOSPHORYLATION SITES.  
RC TISSUE-BRAIN;  
RX MEDLINE; 93135774.  
RA HEEMSKERK F.M., CHEN H.C., HUANG F.L.;  
RT "Protein kinase C phosphorylates Ser-152, Ser-156 and Ser-163 but not  
RT Ser-160 of MARCKS in rat brain".  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:236-241(1993).  
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M59859; -: NOT\_ANNOTATED\_CDS.  
DR PIR; A39773; A39773.  
DR PROSITE; PS00826; MARCKS\_1; 1.  
DR PROSITE; PS00827; MARCKS\_2; 1.  
KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
KW MEMBRANE.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
FT DOMAIN 144 168 CALMODULIN-BINDING (PSD).  
FT MOD\_RES 151 151 PHOSPHORYLATION (BY PKC).  
FT MOD\_RES 155 155 PHOSPHORYLATION (BY PKC).  
FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC).  
SQ SEQUENCE 308 AA; 29663 MW; A6B1CF2A CRC32;  
  
Query Match 84.6%; Score 44; DB 1; Length 308;  
Best Local Similarity 80.0%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KKKKKKFKFK 10  
Db 144 KKKKKRFSFK 153  
|||||:||||  
  
RESULT 6  
MRP\_HUMAN  
ID MRP\_HUMAN STANDARD; PRT; 194 AA.  
AC P49006;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE MARCKS-RELATED PROTEIN (MAC-MARCKS).  
GN MLP OR MRP.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
[1]  
RA BLOCKX H., MAERTENS C., FRANSEN L.M.L.;  
RL SUBMITTED (FEB-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND  
CC CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.  
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X70326; G38435; -;  
DR MIM; 602940; -;  
DR PROSITE; PS00826; MARCKS\_1; 1.  
DR PROSITE; PS00827; MARCKS\_2; 1.  
KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
FT DOMAIN 86 99 CALMODULIN-BINDING (PSD).  
FT MOD\_RES 92 92 POLY-LYS.  
FT MOD\_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
FT MOD\_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
SQ SEQUENCE 194 AA; 19397 MW; D4D367FF CRC32;  
  
Query Match 80.8%; Score 42; DB 1; Length 194;  
Best Local Similarity 88.9%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 KKKKKKFKFK 10  
Db 86 KKKKKRFSFK 94  
|||||:||||  
  
RESULT 7  
MRP\_MOUSE  
ID MRP\_MOUSE STANDARD; PRT; 199 AA.  
AC F28667;  
DT 01-DEC-1992 (REL. 24, CREATED)  
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MARCKS-RELATED PROTEIN (MAC-MARCKS) (BRAIN PROTEIN F52).  
GN MLP OR MRP.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
[1]  
RA STRAIN-BALB/C; TISSUE-BRAIN;  
RX MEDLINE; 91323504.  
RA UMEKAGE T., KATO K.;  
RT "A mouse brain cDNA encodes a novel protein with the protein kinase C  
RT phosphorylation site domain common to MARCKS";  
RL FEBS LETT. 286:147-151(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92386598.  
RA LI J., ADEREM A.;  
RT "MacMARCKS, a novel member of the MARCKS family of protein kinase C  
RT substrates";  
RL CELL 70:791-801(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94010885.  
RA LOBACH D.F., ROCHELLE J.M., WATSON M.L., SELDIN M.F., BLACKSHEAR P.J.;  
RT "Nucleotide sequence, expression, and chromosomal mapping of Mrp and  
RT mapping of five related sequences";  
RL GENOMICS 17:194-204(1993).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE; 92317080.  
RA BLACKSHEAR P.J., VERGHESE G.M., JOHNSON J.D., HAUPT D.M.,  
RA STUMPO D.J.;  
RT "Characteristics of the F52 protein, a MARCKS homologue";  
RL J. BIOL. CHEM. 267:13540-13546(1992).  
CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND  
CC CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.  
CC -!- TISSUE SPECIFICITY: BRAIN (MOSTLY IN DENTATE GYRUS, ANTERIOR  
CC OLFACTORY NUCLEUS, PRIMARY OLFACTORY CORTEX, ENTORHINAL CORTEX,  
CC MEDIAL PREOPTIC AREA, AND DORSOMEDIAL HYPOTHALAMIC NUCLEUS).  
CC

CC -!- PTM: PHOSPHORYLATED.  
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X61399; G50944; -  
CC DR EMBL: S65597; E88783; -  
CC DR PIR: S47185; S17195.  
CC DR PIR: B43341; B43341.  
CC DR MGD: MGI:97143; MLP.  
CC DR PROSITE: PS00826; MARCKS\_1; 1.  
CC DR PROSITE: PS00827; MARCKS\_2; 1.  
CC KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.  
CC FT INIT\_MET 0 0 BY SIMILARITY.  
CC FT LIPID 1 1 MYRISTATE.  
CC FT DOMAIN 86 99 CALMODULIN-BINDING (PSD).  
CC FT MOD\_RES 92 92 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
CC FT MOD\_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
CC FT MOD\_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
CC SEQUENCE 199 AA; 20034 MW; 5CC753C4 CRC32;  
CC -----  
Query Match 80.8%; Score 42; DB 1; Length 199;  
Best Local Similarity 88.9%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KKKKKFKFK 10  
DB 86 KKKKKFSFK 94  
-----  
RESULT 8  
MRP\_RABIT  
ID MRP\_RABIT STANDARD; PRT; 198 AA.  
AC P35566;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MARCKS-RELATED PROTEIN (MAC-MARCKS).  
GN MLP OR MRP.  
OS ORYCTOLAGUS CUNICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE-MACROPHAGE;  
RX MEDLINE: 92386598.  
RA LI J., ADEREM A.;  
RT "MacMarcks", a novel member of the MARCKS family of protein kinase C  
RT substrates.";  
RL CELL 70:791-801(1992).  
CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND  
CC CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.  
CC -!- PTM: PHOSPHORYLATED.  
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: S43921; G255036; -  
CC DR PIR: A43341; A43341.  
CC DR PROSITE: PS00826; MARCKS\_1; 1.

DR PROSITE: PS00827; MARCKS\_2; 1.  
KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
FT DOMAIN 86 99 CALMODULIN-BINDING (PSD).  
FT MOD\_RES 92 92 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
FT MOD\_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
FT MOD\_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
SQ SEQUENCE 198 AA; 19635 MW; 04B10545 CRC32;  
-----  
Query Match 80.8%; Score 42; DB 1; Length 198;  
Best Local Similarity 88.9%; Pred. No. 2.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KKKKKFKFK 10  
DB 86 KKKKKFSFK 94  
-----  
RESULT 9  
VG40\_BPT4  
ID VG40\_BPT4 STANDARD; PRT; 114 AA.  
AC P17171.  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE HEAD FORMATION PROTEIN (PROTEIN GP40) (PROTEIN SP).  
GN 40 OR SP.  
OS BACTERIOPHAGE T4.  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE;  
OC T4-LIKE PHAGES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 89340565.  
RA HINTON D.M.;  
RT "Altered expression of the bacteriophage T4 gene 41  
RT (primase-helicase) in an Escherichia coli rho mutant.";  
RL J. BIOL. CHEM. 264:14440-14446(1989).  
RN [2]  
RP SEQUENCE OF 1-20.  
RX MEDLINE: 89340565.  
RA BLACK L.;  
RT "Altered expression of the bacteriophage T4 gene 41  
RT (primase-helicase) in an Escherichia coli rho mutant.";  
RL UNPUBLISHED RESULTS, CITED BY:  
RL HINTON D.M.;  
RL J. BIOL. CHEM. 264:14440-14446(1989).  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: J04978; G215915; -  
CC DR PIR: B44782; B44782.  
CC SEQUENCE 114 AA; 13290 MW; 8A3CA3A1 CRC32;  
-----  
Query Match 75.0%; Score 39; DB 1; Length 114;  
Best Local Similarity 77.8%; Pred. No. 3.9;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KKKKKFKFK 9  
DB 105 KRSKKFKFK 113  
-----  
RESULT 10  
PAM1\_YEAST

```

ID  PAM1_YEAST      STANDARD;      PRT;      830 AA.
AC  P37304;
DT  01-OCT-1994 (REL. 30, CREATED)
DT  01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT  01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE  PAM1 PROTEIN.
GN  PAM1 OR YDR251W OR YD8419.18 OR YD9320A.01.
OS  SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC  EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCHAROMYCETALES;
OC  SACCHAROMYCETACEAE; SACCHAROMYCES.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=R303-1A.
RX  MEDLINE; 94148839.
RA  HU G.-Z., RONNE H.;
RT  "Overexpression of yeast PAM1 gene permits survival without protein
RT  phosphatase 2A and induces a filamentous phenotype.";
RL  J. BIOL. CHEM. 269:3429-3435(1994).
RN  [2]
RP  SEQUENCE OF 660-830 FROM N.A.
RC  STRAIN=S288C / AB972;
RA  MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL  SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC  -!- FUNCTION: NOT KNOWN. IT IS A SUPPRESSOR OF PROTEIN PHOSPHATASE
CC  2A DEPLETION.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X73454; G456667;
DR  EMBL; Z70202; E228706;
DR  EMBL; Z68329; E215297;
DR  PIR; S45179; S45179.
DR  SGD; L0001334; PAM1.
KW  COILED COIL.
FT  DOMAIN 379 400 COILED COIL.
FT  DOMAIN 481 514 COILED COIL.
FT  DOMAIN 515 522 POLY-GLN.
FT  DOMAIN 810 830 ARG/LYS-RICH (HIGHLY BASIC).
SQ  SEQUENCE 830 AA; 92886 MW; D08F1CD9 CRC32;
-----
Query Match 73.1%; Score 38; DB 1; Length 830;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9
    :|||||
DB 816 EKKKKKFSF 824

RESULT 11
HS9A_CHICK
ID HS9A_CHICK STANDARD; PRT; 728 AA.
AC P11501;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HEAT SHOCK PROTEIN HSP 90-ALPHA.
GN HSPCA.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH, AND GIZZARD;
RX MEDLINE; 89165846.
RA BINART N., CHAMBRAUD B., DUMAS B., ROWLANDS D.A., BIGOGNE C.,

```

```

RA LEVIN J.M., GARNIER J., BAULIEU E.E., CAPELLI M.G.;
RT "The cDNA-derived amino acid sequence of chick heat shock protein Mr
RT 90,000 (HSP 90) reveals a 'DNA like' structure: potential site of
RT interaction with steroid receptors.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 159:140-147(1989).
RN [2]
RP SEQUENCE OF 1-85 FROM N.A.
RX MEDLINE; 89345085.
RA YOURC'H C., BINART N., CHAMBRAUD B., DAVID J.P., JEROME V.,
RA BAILIEU E.E., CAPELLI M.G.;
RT "Isolation and functional analysis of chicken 90-kDa heat shock
RT protein gene promoter.";
RL NUCLEIC ACIDS RES. 17:5259-5272(1989).
CC -!- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; X07265; G63516;
DR  EMBL; X15028; G295722;
DR  PIR; S10880; HCH90.
DR  PROSITE; PS00298; HSP90; 1.
DR  PFAM; PF00183; HSP90; 1.
DR  HSP; P07900; 1YES.
KW  CHAPERONE; ATP-BINDING; HEAT SHOCK; PHOSPHORYLATION.
FT  MOD_RES 230 230 PHOSPHORYLATION (BY SIMILARITY).
FT  MOD_RES 259 259 PHOSPHORYLATION (BY SIMILARITY).
SQ  SEQUENCE 728 AA; 84059 MW; 16C2533A CRC32;
-----
Query Match 71.2%; Score 37; DB 1; Length 728;
Best Local Similarity 80.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10
    :|||||
DB 270 KKKKKKIKK 279

RESULT 12
HS9A_CRIGR
ID HS9A_CRIGR STANDARD; PRT; 732 AA.
AC P46633;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HEAT SHOCK PROTEIN HSP 90-ALPHA (HSP 86).
GN HSPCA OR HSP90A.
OS CRICETULUS GRISEUS (CHINESE HAMSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; CRICETINAE; CRICETULUS.
RN [1]
RP SEQUENCE FROM N.A.
RC CHEN M.S.M.C., LASZLO A.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: L33676; G495759; -.  
 DR PROSITE: PS00298; HSP90; 1.  
 DR PFAM: PF00183; HSP90; 1.  
 DR HSP: P07900; IYES.  
 KW CHAPERONE; ATP-BINDING; HEAT SHOCK; PHOSPHORYLATION.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT MOD\_RES 4 4 PHOSPHORYLATION (BY DS-DNA KINASE)  
 FT (BY SIMILARITY).  
 FT MOD\_RES 6 6 PHOSPHORYLATION (BY DS-DNA KINASE)  
 FT (BY SIMILARITY).  
 FT MOD\_RES 230 230 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 262 262 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 732 AA; 84717 MW; DD0C98E7 CRC32;

Query Match 71.2%; Score 37; DB 1; Length 732;  
 Best Local Similarity 80.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
 DB 274 KKKKKKIKK 283

# RESULT 13

ID HS9A\_HUMAN STANDARD; PRT: 731 AA.  
 AC P07900;  
 DT 01-AUG-1988 (REL. 08, CREATED)  
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HEAT SHOCK PROTEIN HSP 90-ALPHA (HSP 86).  
 GN HSPCA OR HSP90A.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PERIPHERAL BLOOD LYMPHOCYTES;  
 RX MEDLINE: 89386066.  
 RA SOEDA E., YOKOYAMA K., YAMAZAKI M., AKAOGI K., MIWA T., IMAI T.;  
 RT "Nucleotide sequence of a full-length cDNA for 90 kDa heat-shock  
 protein from human peripheral blood lymphocytes.";  
 RL NUCLEIC ACIDS RES. 17:7108-7108(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 91242090.  
 RA YAMAZAKI M., TASHIRO H., YOKOYAMA K., SOEDA E.;  
 RT "Molecular cloning of cDNA encoding a human heat-shock protein whose  
 expression is induced by adenovirus type 12 ELA in HeLa cells.";  
 RL AGRIC. BIOL. CHEM. 54:3163-3170(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PLACENTA;  
 RX MEDLINE: 89343979.  
 RA HICKEY E., BRANDON S.E., SMALE G., LLOYD D., WEBER L.A.;  
 RT "Sequence and regulation of a gene encoding a human 89-kilodalton  
 heat shock protein.";  
 RL MOL. CELL. BIOL. 9:2615-2626(1989).  
 RN [4]  
 RP SEQUENCE OF 1-311 FROM N.A.  
 RX MEDLINE: 88056312.  
 RA HOFFMANN T., HOVMANN B.;  
 RT "Cloning and nucleotide sequence of the murine hsp84 cDNA and  
 chromosome assignment of related sequences.";  
 RL GENE 56:29-40(1987).  
 RN [5]  
 RP SEQUENCE OF 1-311 FROM N.A.  
 RX MEDLINE: 90076956.  
 RA WALTER T., DRABENT B., KREBS H., TOMALAK M., HEISS S.,

RA BENECKE B.J.J.;  
 RT "Cloning and analysis of a human 86-kDa heat-shock-protein-encoding  
 gene.";  
 RL GENE 83:105-115(1989).  
 RN [6]  
 RP SEQUENCE OF 1-20, AND PHOSPHORYLATION.  
 RX MEDLINE: 89123325.  
 RA LEES-MILLER S., ANDERSON C.W.;  
 RT "Two human 90-kDa heat shock proteins are phosphorylated in vivo at  
 conserved serines that are phosphorylated in vitro by casein kinase  
 II.";  
 RL J. BIOL. CHEM. 264:2431-2437(1989).  
 RN [7]  
 RP PHOSPHORYLATION BY DS-DNA KINASE.  
 RX MEDLINE: 9008887.  
 RA LEES-MILLER S., ANDERSON C.W.;  
 RT "The human double-stranded DNA-activated protein kinase phosphorylates  
 the 90-kDa heat-shock protein, hsp90 alpha at two NH2-terminal  
 threonine residues.";  
 RL J. BIOL. CHEM. 264:17275-17280(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 10-222.  
 RX MEDLINE: 97262065.  
 RA STEBBINS C.E., RUSSO A.A., SCHNEIDER C., ROSEN N., HARTL F.U.,  
 RA PAVLETICH N.P.;  
 RT "Crystal structure of an Hsp90-geldanamycin complex: targeting of a  
 protein chaperone by an antitumor agent.";  
 RL CELL 89:239-250(1997).  
 CC -!- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY  
 CC (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X15183; G32488; -.  
 DR EMBL: X07270; G32486; -.  
 DR EMBL: M20724; G703087; -.  
 DR EMBL: M30626; G184419; -.  
 DR PIR: A32319; HHU86.  
 DR PIR: B31420; B31420.  
 DR PIR: JQ0724; JQ0724.  
 DR PDB: 1YER; 22-APR-98.  
 DR PDB: 1YES; 22-APR-98.  
 DR PDB: 1YET; 22-APR-98.  
 DR MIM: 140571; -.  
 DR PROSITE: PS00298; HSP90; 1.  
 DR PFAM: PF00183; HSP90; 1.  
 DR TRANSFAC: T00992; -.  
 KW CHAPERONE; ATP-BINDING; HEAT SHOCK; PHOSPHORYLATION; 3D-STRUCTURE.  
 FT INIT\_MET 0 0  
 FT MOD\_RES 4 4 PHOSPHORYLATION (BY DS-DNA KINASE).  
 FT MOD\_RES 6 6 PHOSPHORYLATION (BY DS-DNA KINASE).  
 FT MOD\_RES 230 230 PHOSPHORYLATION.  
 FT MOD\_RES 262 262 PHOSPHORYLATION.  
 FT CONFLICT 62 62 T -> S (IN REF. 3 TO 5).  
 SQ SEQUENCE 731 AA; 84542 MW; E8152DB4 CRC32;

Query Match 71.2%; Score 37; DB 1; Length 731;  
 Best Local Similarity 80.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
 DB 273 KKKKKKIKK 282



FT MOD\_RES 262 262 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 732 AA; 84643 MW; 3E973712 CRC32;

Query Match 71.2%; Score 37; DB 1; Length 732;  
Best Local Similarity 80.0%; Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KKKKKKFKK 10  
| | | | | | |  
Db 274 KKKKKIKK 283

Search completed: September 7, 1999, 23:50:16  
Job time: 1955 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:27 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-12

Perfect score: 52

Sequence: 1 KKKKKKFK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	84.6	281	A41400	myristylated alani
2	44	84.6	332	A38873	myristylated alani
3	44	84.6	335	S08341	myristylated alani
4	44	84.6	309	A39169	myristylated alani
5	44	84.6	309	A39773	myristylated alani
6	42	80.8	195	S31861	myristylated alani
7	42	80.8	199	A43341	myristylated alani
8	42	80.8	200	S17185	myristylated alani
9	42	80.8	200	A47378	myristylated alani
10	39	75.0	114	B44782	head formation pro
11	38	73.1	830	S54547	PAM1 protein - yea
12	38	73.1	1247	E71616	hypothetical prote
13	38	73.1	374	A42264	membrane-associate
14	37	71.2	732	HHH086	heat shock protein
15	37	71.2	733	HHMS86	heat shock protein
16	37	71.2	728	HHCH90	heat shock protein
17	36	69.2	188	TVHU2K	transforming prote
18	36	69.2	188	S31720	transforming prote
19	36	69.2	188	JCS154	X-ras protein - tu
20	36	69.2	295	H71085	probable MRP prote
21	36	69.2	114	S38226	hypothetical prote
22	36	69.2	204	B71433	hypothetical prote
23	36	69.2	431	T00698	probable translati
24	36	69.2	312	S64436	hypothetical prote
25	36	69.2	1817	D71606	hypothetical prote
26	36	69.2	2013	C71610	probable membrane
27	36	69.2	1516	E71619	RAD2 endonuclease
28	36	69.2	2485	H71621	Ser/Thr protein ki
29	36	69.2	647	A57467	Ralbp1 - rat
30	36	69.2	1116	B70476	hypothetical prote
31	36	69.2	659	E71214	probable V-type so
32	35	67.3	370	S68187	choline-phosphate
33	35	67.3	3951	VFTHB1	F1 protein - avian
34	35	67.3	141	C69007	ribosomal protein
35	35	67.3	999	S15961	hypothetical prote
36	35	67.3	333	JC3229	translation initia
37	35	67.3	333	A31226	translation initia
38	35	67.3	333	S13147	protein synthesis
39	35	67.3	629	A69144	ABC transporter (A

40	35	67.3	396	2	A27891	RepA protein - Bac
41	35	67.3	224	2	C64240	mobilization prote
42	35	67.3	295	2	S34544	hypothetical prote
43	35	67.3	2136	2	A05037	hypothetical prote
44	35	67.3	1051	2	JG4091	glycoprotein A - P
45	35	67.3	2539	2	B71619	hypothetical prote

ALIGNMENTS

RESULT 1

A41400

Myristylated alanine-rich protein kinase C substrate - chicken

C:Species: Gallus gallus (chicken)

C>Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 13-Mar-1998

C:Accession: A41400

R:Graff, J.M.; Stumpo, D.J.; Blackshear, P.J.

Mol. Endocrinol. 3, 1903-1906, 1989

A:Title: Molecular cloning, sequence, and expression of a cDNA encoding the chicken m

A:Reference number: A41400; MUID:90114197

A:Accession: A41400

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-281 <GRA>

A:Cross-references: GB:M31650; NID:g212287; PID:g212288

C:Keywords: actin binding; phosphoprotein

Query Match 84.6%; Score 44; DB 2; Length 281;

Best Local Similarity 80.0%; Pred. No. 2.9;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKFK 10

Db 117 KKKKKRFSK 126

RESULT 2

A38873

Myristylated alanine-rich protein kinase C substrate - human

N:Alternate names: acidic calmodulin-binding 80K protein; MARCKS

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 05-Feb-1999

C:Accession: A38873; A42977; A40758; S29269

R:Shimizu, N

submitted to DBJ, September 1991

A:Reference number: A38873

A:Accession: A38873

A:Molecule type: mRNA

A:Residues: 1-332 <SHI>

A:Cross-references: GB:D10522; GB:D90498; NID:g219893; PID:d1001865; PID:g219894

R:Sakai, K.; Hirai, M.; Kudoh, J.; Minoshima, S.; Shimizu, N.

Genomics 14, 175-178, 1992

A:Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L pro

A:Reference number: A42977; MUID:93052291

A:Accession: A42977

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-13, 'G', 15-332 <SAK>

A:Cross-references: GB:D90498

A:Experimental source: squamous carcinoma cells A431

A:Note: sequence extracted from NCBI backbone (NCBIP:118653)

R:Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.;

J. Biol. Chem. 266, 14399-14405, 1991

A:Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACKS)

A:Reference number: A40758; MUID:91317795

A:Accession: A40758

A:Molecule type: mRNA

A:Residues: 1-83, 'A', 85-118, 'P', 120-233, 'W', 235-286, 'LVC', 290, 'RRGSPRGARGRRSLNQ', 30

A:Cross-references: GB:M68956

A:Note: the authors translated the codon GGC for residue 53 as Arg

R:Herget, T.; Brooks, S.F.; Broad, S.; Rozenfurt, E.

Eur. J. Biochem. 209, 7-14, 1992  
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein or equivalent genes in different species.  
A:Reference number: S29267; MUID:93011168  
A:Accession: S29269  
A:Molecule type: mRNA  
A:Residues: 189-223, R, 225-234, 'E', 236-322 <HER>  
C:Comment: This protein is a major cellular substrate for protein kinase C and plays a role in the presence of calcium and calmodulin.  
C:Genetics:  
A:Gene: GDB:WACS  
A:CROSS-references: GDB:118835; OMIM:177061  
A:Map position: 6q22.2-6q22.2  
C:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristylation; myristoylated amino end (Gly) (in mature form) #status predicted  
F:159,163,167,170/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 84.6%; Score 44; DB 2; Length 332;  
Best Local Similarity 80.0%; Pred. No. 3.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10  
|||||:|  
Db 152 KKKKKRFSFK 161

RESULT 3  
S08341  
myristoylated alanine-rich protein kinase C substrate - bovine  
N:Alternate names: acidic calmodulin-binding 81K protein (ACAMP-81); MARCKS  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 05-Feb-1999  
C:Accession: S08341; A32904; S29270; A46098; P50338  
R:Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.  
Nucleic Acids Res. 17, 3987-3988, 1989  
A:Title: Nucleotide sequence of a cDNA for the bovine myristoylated alanine-rich C kinase  
A:Reference number: S08341; MUID:89282412  
A:Accession: S08341  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-335 <STU>  
A:CROSS-references: EMBL:M24638; NID:g163339; PID:g163340  
R:Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4012-4016, 1989  
A:Title: Molecular cloning, characterization, and expression of a cDNA encoding the "80-kDa"  
A:Reference number: A32904; MUID:89264553  
A:Accession: A32904  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-98, 'Q', 100-335 <ST2>  
A:CROSS-references: GB:M24638; GB:M23738  
R:Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
Eur. J. Biochem. 209, 7-14, 1992  
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein or equivalent genes in different species.  
A:Reference number: S29267; MUID:93011168  
A:Accession: S29270  
A:Molecule type: DNA  
A:Residues: 191-253, 'SEE', 257-279, 283-292, 'V', 294, 'PQEE', 299, 'A', 302-313, 'A', 315-319, 'A', 320-335, 'A', 336-338, 'A', 339-340, 'A', 341-342, 'A', 343-344, 'A', 345-346, 'A', 347-348, 'A', 349-350, 'A', 351-352, 'A', 353-354, 'A', 355-356, 'A', 357-358, 'A', 359-360, 'A', 361-362, 'A', 363-364, 'A', 365-366, 'A', 367-368, 'A', 369-370, 'A', 371-372, 'A', 373-374, 'A', 375-376, 'A', 377-378, 'A', 379-380, 'A', 381-382, 'A', 383-384, 'A', 385-386, 'A', 387-388, 'A', 389-390, 'A', 391-392, 'A', 393-394, 'A', 395-396, 'A', 397-398, 'A', 399-400, 'A', 401-402, 'A', 403-404, 'A', 405-406, 'A', 407-408, 'A', 409-410, 'A', 411-412, 'A', 413-414, 'A', 415-416, 'A', 417-418, 'A', 419-420, 'A', 421-422, 'A', 423-424, 'A', 425-426, 'A', 427-428, 'A', 429-430, 'A', 431-432, 'A', 433-434, 'A', 435-436, 'A', 437-438, 'A', 439-440, 'A', 441-442, 'A', 443-444, 'A', 445-446, 'A', 447-448, 'A', 449-450, 'A', 451-452, 'A', 453-454, 'A', 455-456, 'A', 457-458, 'A', 459-460, 'A', 461-462, 'A', 463-464, 'A', 465-466, 'A', 467-468, 'A', 469-470, 'A', 471-472, 'A', 473-474, 'A', 475-476, 'A', 477-478, 'A', 479-480, 'A', 481-482, 'A', 483-484, 'A', 485-486, 'A', 487-488, 'A', 489-490, 'A', 491-492, 'A', 493-494, 'A', 495-496, 'A', 497-498, 'A', 499-500, 'A', 501-502, 'A', 503-504, 'A', 505-506, 'A', 507-508, 'A', 509-510, 'A', 511-512, 'A', 513-514, 'A', 515-516, 'A', 517-518, 'A', 519-520, 'A', 521-522, 'A', 523-524, 'A', 525-526, 'A', 527-528, 'A', 529-530, 'A', 531-532, 'A', 533-534, 'A', 535-536, 'A', 537-538, 'A', 539-540, 'A', 541-542, 'A', 543-544, 'A', 545-546, 'A', 547-548, 'A', 549-550, 'A', 551-552, 'A', 553-554, 'A', 555-556, 'A', 557-558, 'A', 559-560, 'A', 561-562, 'A', 563-564, 'A', 565-566, 'A', 567-568, 'A', 569-570, 'A', 571-572, 'A', 573-574, 'A', 575-576, 'A', 577-578, 'A', 579-580, 'A', 581-582, 'A', 583-584, 'A', 585-586, 'A', 587-588, 'A', 589-590, 'A', 591-592, 'A', 593-594, 'A', 595-596, 'A', 597-598, 'A', 599-600, 'A', 601-602, 'A', 603-604, 'A', 605-606, 'A', 607-608, 'A', 609-610, 'A', 611-612, 'A', 613-614, 'A', 615-616, 'A', 617-618, 'A', 619-620, 'A', 621-622, 'A', 623-624, 'A', 625-626, 'A', 627-628, 'A', 629-630, 'A', 631-632, 'A', 633-634, 'A', 635-636, 'A', 637-638, 'A', 639-640, 'A', 641-642, 'A', 643-644, 'A', 645-646, 'A', 647-648, 'A', 649-650, 'A', 651-652, 'A', 653-654, 'A', 655-656, 'A', 657-658, 'A', 659-660, 'A', 661-662, 'A', 663-664, 'A', 665-666, 'A', 667-668, 'A', 669-670, 'A', 671-672, 'A', 673-674, 'A', 675-676, 'A', 677-678, 'A', 679-680, 'A', 681-682, 'A', 683-684, 'A', 685-686, 'A', 687-688, 'A', 689-690, 'A', 691-692, 'A', 693-694, 'A', 695-696, 'A', 697-698, 'A', 699-700, 'A', 701-702, 'A', 703-704, 'A', 705-706, 'A', 707-708, 'A', 709-710, 'A', 711-712, 'A', 713-714, 'A', 715-716, 'A', 717-718, 'A', 719-720, 'A', 721-722, 'A', 723-724, 'A', 725-726, 'A', 727-728, 'A', 729-730, 'A', 731-732, 'A', 733-734, 'A', 735-736, 'A', 737-738, 'A', 739-740, 'A', 741-742, 'A', 743-744, 'A', 745-746, 'A', 747-748, 'A', 749-750, 'A', 751-752, 'A', 753-754, 'A', 755-756, 'A', 757-758, 'A', 759-760, 'A', 761-762, 'A', 763-764, 'A', 765-766, 'A', 767-768, 'A', 769-770, 'A', 771-772, 'A', 773-774, 'A', 775-776, 'A', 777-778, 'A', 779-780, 'A', 781-782, 'A', 783-784, 'A', 785-786, 'A', 787-788, 'A', 789-790, 'A', 791-792, 'A', 793-794, 'A', 795-796, 'A', 797-798, 'A', 799-800, 'A', 801-802, 'A', 803-804, 'A', 805-806, 'A', 807-808, 'A', 809-810, 'A', 811-812, 'A', 813-814, 'A', 815-816, 'A', 817-818, 'A', 819-820, 'A', 821-822, 'A', 823-824, 'A', 825-826, 'A', 827-828, 'A', 829-830, 'A', 831-832, 'A', 833-834, 'A', 835-836, 'A', 837-838, 'A', 839-840, 'A', 841-842, 'A', 843-844, 'A', 845-846, 'A', 847-848, 'A', 849-850, 'A', 851-852, 'A', 853-854, 'A', 855-856, 'A', 857-858, 'A', 859-860, 'A', 861-862, 'A', 863-864, 'A', 865-866, 'A', 867-868, 'A', 869-870, 'A', 871-872, 'A', 873-874, 'A', 875-876, 'A', 877-878, 'A', 879-880, 'A', 881-882, 'A', 883-884, 'A', 885-886, 'A', 887-888, 'A', 889-890, 'A', 891-892, 'A', 893-894, 'A', 895-896, 'A', 897-898, 'A', 899-900, 'A', 901-902, 'A', 903-904, 'A', 905-906, 'A', 907-908, 'A', 909-910, 'A', 911-912, 'A', 913-914, 'A', 915-916, 'A', 917-918, 'A', 919-920, 'A', 921-922, 'A', 923-924, 'A', 925-926, 'A', 927-928, 'A', 929-930, 'A', 931-932, 'A', 933-934, 'A', 935-936, 'A', 937-938, 'A', 939-940, 'A', 941-942, 'A', 943-944, 'A', 945-946, 'A', 947-948, 'A', 949-950, 'A', 951-952, 'A', 953-954, 'A', 955-956, 'A', 957-958, 'A', 959-960, 'A', 961-962, 'A', 963-964, 'A', 965-966, 'A', 967-968, 'A', 969-970, 'A', 971-972, 'A', 973-974, 'A', 975-976, 'A', 977-978, 'A', 979-980, 'A', 981-982, 'A', 983-984, 'A', 985-986, 'A', 987-988, 'A', 989-990, 'A', 991-992, 'A', 993-994, 'A', 995-996, 'A', 997-998, 'A', 999-1000, 'A', 1001-1002, 'A', 1003-1004, 'A', 1005-1006, 'A', 1007-1008, 'A', 1009-1010, 'A', 1011-1012, 'A', 1013-1014, 'A', 1015-1016, 'A', 1017-1018, 'A', 1019-1020, 'A', 1021-1022, 'A', 1023-1024, 'A', 1025-1026, 'A', 1027-1028, 'A', 1029-1030, 'A', 1031-1032, 'A', 1033-1034, 'A', 1035-1036, 'A', 1037-1038, 'A', 1039-1040, 'A', 1041-1042, 'A', 1043-1044, 'A', 1045-1046, 'A', 1047-1048, 'A', 1049-1050, 'A', 1051-1052, 'A', 1053-1054, 'A', 1055-1056, 'A', 1057-1058, 'A', 1059-1060, 'A', 1061-1062, 'A', 1063-1064, 'A', 1065-1066, 'A', 1067-1068, 'A', 1069-1070, 'A', 1071-1072, 'A', 1073-1074, 'A', 1075-1076, 'A', 1077-1078, 'A', 1079-1080, 'A', 1081-1082, 'A', 1083-1084, 'A', 1085-1086, 'A', 1087-1088, 'A', 1089-1090, 'A', 1091-1092, 'A', 1093-1094, 'A', 1095-1096, 'A', 1097-1098, 'A', 1099-1100, 'A', 1101-1102, 'A', 1103-1104, 'A', 1105-1106, 'A', 1107-1108, 'A', 1109-1110, 'A', 1111-1112, 'A', 1113-1114, 'A', 1115-1116, 'A', 1117-1118, 'A', 1119-1120, 'A', 1121-1122, 'A', 1123-1124, 'A', 1125-1126, 'A', 1127-1128, 'A', 1129-1130, 'A', 1131-1132, 'A', 1133-1134, 'A', 1135-1136, 'A', 1137-1138, 'A', 1139-1140, 'A', 1141-1142, 'A', 1143-1144, 'A', 1145-1146, 'A', 1147-1148, 'A', 1149-1150, 'A', 1151-1152, 'A', 1153-1154, 'A', 1155-1156, 'A', 1157-1158, 'A', 1159-1160, 'A', 1161-1162, 'A', 1163-1164, 'A', 1165-1166, 'A', 1167-1168, 'A', 1169-1170, 'A', 1171-1172, 'A', 1173-1174, 'A', 1175-1176, 'A', 1177-1178, 'A', 1179-1180, 'A', 1181-1182, 'A', 1183-1184, 'A', 1185-1186, 'A', 1187-1188, 'A', 1189-1190, 'A', 1191-1192, 'A', 1193-1194, 'A', 1195-1196, 'A', 1197-1198, 'A', 1199-1200, 'A', 1201-1202, 'A', 1203-1204, 'A', 1205-1206, 'A', 1207-1208, 'A', 1209-1210, 'A', 1211-1212, 'A', 1213-1214, 'A', 1215-1216, 'A', 1217-1218, 'A', 1219-1220, 'A', 1221-1222, 'A', 1223-1224, 'A', 1225-1226, 'A', 1227-1228, 'A', 1229-1230, 'A', 1231-1232, 'A', 1233-1234, 'A', 1235-1236, 'A', 1237-1238, 'A', 1239-1240, 'A', 1241-1242, 'A', 1243-1244, 'A', 1245-1246, 'A', 1247-1248, 'A', 1249-1250, 'A', 1251-1252, 'A', 1253-1254, 'A', 1255-1256, 'A', 1257-1258, 'A', 1259-1260, 'A', 1261-1262, 'A', 1263-1264, 'A', 1265-1266, 'A', 1267-1268, 'A', 1269-1270, 'A', 1271-1272, 'A', 1273-1274, 'A', 1275-1276, 'A', 1277-1278, 'A', 1279-1280, 'A', 1281-1282, 'A', 1283-1284, 'A', 1285-1286, 'A', 1287-1288, 'A', 1289-1290, 'A', 1291-1292, 'A', 1293-1294, 'A', 1295-1296, 'A', 1297-1298, 'A', 1299-1300, 'A', 1301-1302, 'A', 1303-1304, 'A', 1305-1306, 'A', 1307-1308, 'A', 1309-1310, 'A', 1311-1312, 'A', 1313-1314, 'A', 1315-1316, 'A', 1317-1318, 'A', 1319-1320, 'A', 1321-1322, 'A', 1323-1324, 'A', 1325-1326, 'A', 1327-1328, 'A', 1329-1330, 'A', 1331-1332, 'A', 1333-1334, 'A', 1335-1336, 'A', 1337-1338, 'A', 1339-1340, 'A', 1341-1342, 'A', 1343-1344, 'A', 1345-1346, 'A', 1347-1348, 'A', 1349-1350, 'A', 1351-1352, 'A', 1353-1354, 'A', 1355-1356, 'A', 1357-1358, 'A', 1359-1360, 'A', 1361-1362, 'A', 1363-1364, 'A', 1365-1366, 'A', 1367-1368, 'A', 1369-1370, 'A', 1371-1372, 'A', 1373-1374, 'A', 1375-1376, 'A', 1377-1378, 'A', 1379-1380, 'A', 1381-1382, 'A', 1383-1384, 'A', 1385-1386, 'A', 1387-1388, 'A', 1389-1390, 'A', 1391-1392, 'A', 1393-1394, 'A', 1395-1396, 'A', 1397-1398, 'A', 1399-1400, 'A', 1401-1402, 'A', 1403-1404, 'A', 1405-1406, 'A', 1407-1408, 'A', 1409-1410, 'A', 1411-1412, 'A', 1413-1414, 'A', 1415-1416, 'A', 1417-1418, 'A', 1419-1420, 'A', 1421-1422, 'A', 1423-1424, 'A', 1425-1426, 'A', 1427-1428, 'A', 1429-1430, 'A', 1431-1432, 'A', 1433-1434, 'A', 1435-1436, 'A', 1437-1438, 'A', 1439-1440, 'A', 1441-1442, 'A', 1443-1444, 'A', 1445-1446, 'A', 1447-1448, 'A', 1449-1450, 'A', 1451-1452, 'A', 1453-1454, 'A', 1455-1456, 'A', 1457-1458, 'A', 1459-1460, 'A', 1461-1462, 'A', 1463-1464, 'A', 1465-1466, 'A', 1467-1468, 'A', 1469-1470, 'A', 1471-1472, 'A', 1473-1474, 'A', 1475-1476, 'A', 1477-1478, 'A', 1479-1480, 'A', 1481-1482, 'A', 1483-1484, 'A', 1485-1486, 'A', 1487-1488, 'A', 1489-1490, 'A', 1491-1492, 'A', 1493-1494, 'A', 1495-1496, 'A', 1497-1498, 'A', 1499-1500, 'A', 1501-1502, 'A', 1503-1504, 'A', 1505-1506, 'A', 1507-1508, 'A', 1509-1510, 'A', 1511-1512, 'A', 1513-1514, 'A', 1515-1516, 'A', 1517-1518, 'A', 1519-1520, 'A', 1521-1522, 'A', 1523-1524, 'A', 1525-1526, 'A', 1527-1528, 'A', 1529-1530, 'A', 1531-1532, 'A', 1533-1534, 'A', 1535-1536, 'A', 1537-1538, 'A', 1539-1540, 'A', 1541-1542, 'A', 1543-1544, 'A', 1545-1546, 'A', 1547-1548, 'A', 1549-1550, 'A', 1551-1552, 'A', 1553-1554, 'A', 1555-1556, 'A', 1557-1558, 'A', 1559-1560, 'A', 1561-1562, 'A', 1563-1564, 'A', 1565-1566, 'A', 1567-1568, 'A', 1569-1570, 'A', 1571-1572, 'A', 1573-1574, 'A', 1575-1576, 'A', 1577-1578, 'A', 1579-1580, 'A', 1581-1582, 'A', 1583-1584, 'A', 1585-1586, 'A', 1587-1588, 'A', 1589-1590, 'A', 1591-1592, 'A', 1593-1594, 'A', 1595-1596, 'A', 1597-1598, 'A', 1599-1600, 'A', 1601-1602, 'A', 1603-1604, 'A', 1605-1606, 'A', 1607-1608, 'A', 1609-1610, 'A', 1611-1612, 'A', 1613-1614, 'A', 1615-1616, 'A', 1617-1618, 'A', 1619-1620, 'A', 1621-1622, 'A', 1623-1624, 'A', 1625-1626, 'A', 1627-1628, 'A', 1629-1630, 'A', 1631-1632, 'A', 1633-1634, 'A', 1635-1636, 'A', 1637-1638, 'A', 1639-1640, 'A', 1641-1642, 'A', 1643-1644, 'A', 1645-1646, 'A', 1647-1648, 'A', 1649-1650, 'A', 1651-1652, 'A', 1653-1654, 'A', 1655-1656, 'A', 1657-1658, 'A', 1659-1660, 'A', 1661-1662, 'A', 1663-1664, 'A', 1665-1666, 'A', 1667-1668, 'A', 1669-1670, 'A', 1671-1672, 'A', 1673-1674, 'A', 1675-1676, 'A', 1677-1678, 'A', 1679-1680, 'A', 1681-1682, 'A', 1683-1684, 'A', 1685-1686, 'A', 1687-1688, 'A', 1689-1690, 'A', 1691-1692, 'A', 1693-1694, 'A', 1695-1696, 'A', 1697-1698, 'A', 1699-1700, 'A', 1701-1702, 'A', 1703-1704, 'A', 1705-1706, 'A', 1707-1708, 'A', 1709-1710, 'A', 1711-1712, 'A', 1713-1714, 'A', 1715-1716, 'A', 1717-1718, 'A', 1719-1720, 'A', 1721-1722, 'A', 1723-1724, 'A', 1725-1726, 'A', 1727-1728, 'A', 1729-1730, 'A', 1731-1732, 'A', 1733-1734, 'A', 1735-1736, 'A', 1737-1738, 'A', 1739-1740, 'A', 1741-1742, 'A', 1743-1744, 'A', 1745-1746, 'A', 1747-1748, 'A', 1749-1750, 'A', 1751-1752, 'A', 1753-1754, 'A', 1755-1756, 'A', 1757-1758, 'A', 1759-1760, 'A', 1761-1762, 'A', 1763-1764, 'A', 1765-1766, 'A', 1767-1768, 'A', 1769-1770, 'A', 1771-1772, 'A', 1773-1774, 'A', 1775-1776, 'A', 1777-1778, 'A', 1779-1780, 'A', 1781-1782, 'A', 1783-1784, 'A', 1785-1786, 'A', 1787-1788, 'A', 1789-1790, 'A', 1791-1792, 'A', 1793-1794, 'A', 1795-1796, 'A', 1797-1798, 'A', 1799-1800, 'A', 1801-1802, 'A', 1803-1804, 'A', 1805-1806, 'A', 1807-1808, 'A', 1809-1810, 'A', 1811-1812, 'A', 1813-1814, 'A', 1815-1816, 'A', 1817-1818, 'A', 1819-1820, 'A', 1821-1822, 'A', 1823-1824, 'A', 1825-1826, 'A', 1827-1828, 'A', 1829-1830, 'A', 1831-1832, 'A', 1833-1834, 'A', 1835-1836, 'A', 1837-1838, 'A', 1839-1840, 'A', 1841-1842, 'A', 1843-1844, 'A', 1845-1846, 'A', 1847-1848, 'A', 1849-1850, 'A', 1851-1852, 'A', 1853-1854, 'A', 1855-1856, 'A', 1857-1858, 'A', 1859-1860, 'A', 1861-1862, 'A', 1863-1864, 'A', 1865-1866, 'A', 1867-1868, 'A', 1869-1870, 'A', 1871-1872, 'A', 1873-1874, 'A', 1875-1876, 'A', 1877-1878, 'A', 1879-1880, 'A', 1881-1882, 'A', 1883-1884, 'A', 1885-1886, 'A', 1887-1888, 'A', 1889-1890, 'A', 1891-1892, 'A', 1893-1894, 'A', 1895-1896, 'A', 1897-1898, 'A', 1899-1900, 'A', 1901-1902, 'A', 1903-1904, 'A', 1905-1906, 'A', 1907-1908, 'A', 1909-1910, 'A', 1911-1912, 'A', 1913-1914, 'A', 1915-1916, 'A', 1917-1918, 'A', 1919-1920, 'A', 1921-1922, 'A', 1923-1924, 'A', 1925-1926, 'A', 1927-1928, 'A', 1929-1930, 'A', 1931-1932, 'A', 1933-1934, 'A', 1935-1936, 'A', 1937-1938, 'A', 1939-1940, 'A', 1941-1942, 'A', 1943-1944, 'A', 1945-1946, 'A', 1947-1948, 'A', 1949-1950, 'A', 1951-1952, 'A', 1953-1954, 'A', 1955-1956, 'A', 1957-1958, 'A', 1959-1960, 'A', 1961-1962, 'A', 1963-1964, 'A', 1965-1966, 'A', 1967-1968, 'A', 1969-1970, 'A', 1971-1972, 'A', 1973-1974, 'A', 1975-1976, 'A', 1977-1978, 'A', 1979-1980, 'A', 1981-1982, 'A', 1983-1984, 'A', 1985-1986, 'A', 1987-1988, 'A', 1989-1990, 'A', 1991-1992, 'A', 1993-1994, 'A', 1995-1996,

Best Local Similarity 80.0%; Pred. No. 3.1; Mismatches 1; Indels 0; Gaps 0;  
Matches 8; Conservative 1;  
QY 1 KKKKKKFKFK 10  
|||||:|  
Db 145 KKKKKRFSFK 154

RESULT 5  
A39773  
myristylated alanine-rich protein kinase C substrate, macrophage - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 13-Mar-1998  
C:Accession: A39773  
R:Brusilinsky, J.D.; Brooks, S.F.; Herget, T.; Morris, C.; Rozenfurt, E.  
J. Biol. Chem. 266, 7073-7080, 1991  
A:Title: Molecular cloning and characterization of the acidic 80-kDa protein kinase C sub-  
A:Reference number: A39773; MUID:91201362  
A:Accession: A39773  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-309 <ERU>  
A:Cross-references: GB:M59859  
C:Keywords: phosphoprotein

Query Match 84.6%; Score 44; DB 2; Length 309;  
Best Local Similarity 80.0%; Pred. No. 3.1; Mismatches 1; Indels 0; Gaps 0;  
Matches 8; Conservative 1;  
QY 1 KKKKKKFKFK 10  
|||||:|  
Db 145 KKKKKRFSFK 154

RESULT 6  
S31861  
myristylated alanine-rich protein kinase C substrate, macrophage - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Sep-1997  
C:Accession: S31861  
R:Blockx, H.; Maertens, C.; Fransen, L.M.L.  
submitted to the EMBL Data Library, February 1993  
A:Description: cDNA and derived amino acid sequence of human maemarcks an LPS-inducible  
A:Reference number: S31861  
A:Accession: S31861  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-195 <BLO>  
A:Cross-references: EMBL:X70326; NID:g38434; PID:g38435

Query Match 80.8%; Score 42; DB 2; Length 195;  
Best Local Similarity 88.9%; Pred. No. 4.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0;  
QY 2 KKKKKKFKFK 10  
|||||:|  
Db 87 KKKKKRFSFK 95

RESULT 7  
A43341  
myristylated alanine-rich protein kinase C substrate, macrophage - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
C:Accession: A43341  
R:Li, J.; Aderem, A.  
Cell 70, 791-801, 1992  
A:Title: MacMARCKS, a novel member of the MARCKS family of protein kinase C substrates.  
A:Reference number: A43341; MUID:92386598  
A:Accession: A43341  
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid; protein  
A:Residues: 1-199 <LII>  
A:Cross-references: GB:S43921; NID:g255035; PID:g255036  
A:Experimental source: alveolar macrophage  
A:Note: sequence extracted from NCBI backbone (NCBIP:112773)  
C:Keywords: phosphoprotein

Query Match 80.8%; Score 42; DB 2; Length 199;  
Best Local Similarity 88.9%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0;  
QY 2 KKKKKKFKFK 10  
|||||:|  
Db 87 KKKKKRFSFK 95

RESULT 8  
S17185  
myristylated alanine-rich protein kinase C substrate - mouse  
N:Alternate names: MacMARCKS  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 13-Mar-1998  
C:Accession: S17185; B43341  
R:Umekage, T.; Kato, K.  
FEBS Lett. 286, 147-151, 1991  
A:Title: A mouse brain cDNA encodes a novel protein with the protein kinase C phospho  
A:Reference number: S17185; MUID:91323504  
A:Accession: S17185  
A:Molecule type: mRNA  
A:Residues: 1-200 <DME>  
A:Cross-references: EMBL:X61399; NID:g50943; PID:g50944  
R:Li, J.; Aderem, A.  
Cell 70, 791-801, 1992  
A:Title: MacMARCKS, a novel member of the MARCKS family of protein kinase C substrate  
A:Reference number: A43341; MUID:92386598  
A:Accession: B43341  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-200 <LII>  
A:Experimental source: resident peritoneal macrophage  
A:Note: sequence extracted from NCBI backbone (NCBIP:112774)  
C:Keywords: phosphoprotein

Query Match 80.8%; Score 42; DB 2; Length 200;  
Best Local Similarity 88.9%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0;  
QY 2 KKKKKKFKFK 10  
|||||:|  
Db 87 KKKKKRFSFK 95

RESULT 9  
A47378  
myristylated alanine-rich protein kinase C substrate-related protein - mouse  
N:Alternate names: Mrp protein  
C:Species: Mus sp. (mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 17-Mar-1999  
C:Accession: A47378  
R:Lobach, D.F.; Rochelle, J.M.; Watson, M.L.; Seidlin, M.F.; Blackshear, P.J.  
Genomics 17, 194-204, 1993  
A:Title: Nucleotide sequence, expression, and chromosomal mapping of Mrp and mapping  
A:Reference number: A47378; MUID:94010885  
A:Accession: A47378  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-200 <RES>  
A:Cross-references: GB:S65597; NID:9415565  
C:Genetics:  
A:Gene: Mrp  
A:Introns: 29/3

Query Match 80.8%; Score 42; DB 2; Length 200;  
Best Local Similarity 88.9%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKKKFK 10  
||||| ||  
DB 87 KKKKKKFSK 95

RESULT 10  
B44782  
head formation protein - phage T4  
N:Alternate names: gp 40  
C:Species: phage T4  
A:Note: host Escherichia coli  
C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Sep-1997  
C:Accession: B44782; J20013  
R:Hinton, D.M.  
J. Biol. Chem. 264, 14440-14446, 1989  
A:Title: Altered expression of the bacteriophage T4 gene 41 (primase-helicase) in an Es  
A:Reference number: A44782; MUID:89340565  
A:Accession: B44782  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-114 <HIN>  
A:Cross-references: GB:J04978; NID:g215913; PID:g215915  
C:Genetics:  
A:Gene: 40  
A:Map position: 22.056-22.398

Query Match 75.0%; Score 39; DB 2; Length 114;  
Best Local Similarity 77.8%; Pred. No. 7.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9  
|:|||||  
DB 105 KRSKKKFKF 113

RESULT 11  
S54547  
PAM1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YD8419.18; protein YD93320A.01; protein YDR251W  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 23-Aug-1996 #text\_change 06-Feb-1998  
C:Accession: S54547; A53049; S67454; S45179  
R:Oliver, K.; Harris, D.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54530  
A:Accession: S54547  
A:Molecule type: DNA  
A:Residues: 1-659 <OLI>  
A:Cross-references: EMBL:249701; NID:g817819; PID:g817837; MIPS:YDR251W  
R:Hu, G.Z.; Ronne, H.  
J. Biol. Chem. 269, 3429-3435, 1994  
A:Title: Overexpression of yeast PAM1 gene permits survival without protein phosphatase  
A:Reference number: A53049; MUID:94148839  
A:Accession: A53049  
A:Molecule type: DNA  
A:Residues: 1-255, 'T', 257-830 <HUA>  
A:Cross-references: EMBL:X73454; NID:g456666; PID:g456667  
R:Murphy, L.; Harris, D.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: S61117  
A:Accession: S67308  
A:Molecule type: DNA  
A:Residues: 660-830 <MUR>  
A:Cross-references: EMBL:Z668329; NID:g1136205; PID:e215297; PID:g1136206; MIPS:YDR251W  
R:Murphy, L.; Harris, D.  
submitted to the EMBL Data Library, March 1996

A:Reference number: S67454  
A:Accession: S67454  
A:Molecule type: DNA  
A:Residues: 660-830 <MUR>  
A:Cross-references: EMBL:Z70202; NID:g1226026; PID:e228706; PID:g1226027  
C:Genetics:  
A:Gene: SGD:PAM1  
A:Cross-references: SGD:S0002659; MIPS:YDR251W  
A:Map position: 4R  
C:Keywords: coiled coil

Query Match 73.1%; Score 38; DB 2; Length 830;  
Best Local Similarity 77.8%; Pred. No. 63;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9  
:||||| |  
DB 816 ERKKKKFSF 824

RESULT 12  
E71616  
hypothetical protein PFB0365w - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: E71616  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
; Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: E71616  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1247 <GAR>  
A:Cross-references: GB:AE001390; GB:AE001362; NID:g3845164; PID:g3845165; TIGR:PFB036  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0365w

Query Match 73.1%; Score 38; DB 2; Length 1247;  
Best Local Similarity 87.5%; Pred. No. 90;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFK 8  
|:|||||  
DB 483 KKKKKKFK 490

RESULT 13  
A42264  
membrane-associated calcium-binding protein Pfs40 - Plasmodium falciparum  
C:Species: Plasmodium falciparum  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: A42264  
R:Rawlings, D.J.; Kaslow, D.C.  
J. Biol. Chem. 267, 3976-3982, 1992  
A:Title: A novel 40-kDa membrane-associated EF-hand calcium-binding protein in Plasm  
A:Reference number: A42264; MUID:92156141  
A:Accession: A42264  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid; protein  
A:Residues: 1-374 <RAW>  
A:Note: sequence extracted from NCBI backbone (NCBIP:82671)  
C:Keywords: calcium binding

Query Match 73.1%; Score 38; DB 2; Length 374;  
Best Local Similarity 87.5%; Pred. No. 30;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 1 KKKKKKFK 8  
HHU86  
Db 363 KKKKKKFR 370

## RESULT 14

heat shock protein 90-alpha - human  
N:Alternate names: heat shock protein 86; heat shock protein 89 alpha  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 20-Mar-1998  
C:Accession: A32319; J00128; PS0020; S06898; J00724; B31420; S67961  
R:Hickey, E.; Brandon, S.E.; Smale, G.; Lloyd, D.; Weber, L.A.  
Mol. Cell. Biol. 9, 2615-2626, 1989  
A:Title: Sequence and regulation of a gene encoding a human 89-kilodalton heat shock protein  
A:Reference number: A32319; MUID:89343979  
A:Accession: A32319  
A:Molecule type: DNA  
A:Residues: 1-732 <HIC>  
A:Cross-references: GB:M27024; NID:g341598; PID:g703087  
A:Note: the authors translated the codon AAC for residue 383 as Asp  
R:Walter, T.; Drabent, B.; Krebs, H.; Tomalak, M.; Helss, S.; Benecke, B.J.  
Gene 83, 105-115, 1989  
A:Title: Cloning and analysis of a human 86-kDa heat-shock-protein-encoding gene.  
A:Reference number: JQ0128; MUID:90076956  
A:Accession: JQ0128  
A:Molecule type: DNA  
A:Residues: 1-312 <WAL>  
A:Cross-references: GB:M30626; NID:g184418; PID:g184419  
A:Note: the authors translated the codon AAA for residue 58 as Leu and AAC for residue 8  
R:Hoffmann, T.; Hovemann, B.  
Gene 74, 491-501, 1988  
A:Title: Heat-shock proteins, Hsp84 and Hsp86, of mice and men: two related genes encoded  
A:Reference number: PS0020; MUID:89232740  
A:Accession: PS0020  
A:Molecule type: mRNA  
A:Residues: 1-312 <HOF>  
A:Cross-references: GB:X07270; NID:g32485; PID:g32486  
R:Yamazaki, M.; Akaogi, K.; Miwa, T.; Inai, T.; Soeda, E.; Yokoyama, K.  
Nucleic Acids Res. 17, 7108, 1989  
A:Title: Nucleotide sequence of a full-length cDNA for 90 kDa heat-shock protein from hu  
A:Reference number: S06898; MUID:89386066  
A:Accession: S06898  
A:Molecule type: mRNA  
A:Residues: 1-20, 'LSG', 25-62, 'T', 64-161, 'GVLSR', 168-732 <YAM>  
A:Experimental source: peripheral blood lymphocytes  
R:Yamazaki, M.; Tashiro, H.; Yokoyama, K.; Soeda, E.  
Agric. Biol. Chem. 54, 3163-3170, 1990  
A:Title: Molecular cloning of cDNA encoding a human heat-shock protein whose expression  
A:Reference number: JQ0724  
A:Accession: JQ0724  
A:Molecule type: mRNA  
A:Residues: 1-62, 'T', 64-732 <YA2>  
A:Experimental source: adenovirus type 12 ELA-transfected HeLa cells  
R:Lees-Miller, S.P.; Anderson, C.W.  
J. Biol. Chem. 264, 2431-2437, 1989  
A:Title: Two human 90-kDa heat shock proteins are phosphorylated in vivo at conserved se  
A:Reference number: A92741; MUID:89123325  
A:Accession: B31420  
A:Molecule type: protein  
A:Residues: 2-21;223-274 <LEE>  
A:Experimental source: HeLa cells in exponential growth  
A:Note: phosphorylation sites were determined by metabolic labelling with [32-P]orthophos  
R:Nemoto, T.; Ohara-Nemoto, Y.; Ota, M.; Takagi, T.; Yokoyama, K.  
Eur. J. Biochem. 233, 1-8, 1995  
A:Title: Mechanism of dimer formation of the 90-kDa heat-shock protein.  
A:Reference number: S67961; MUID:96061925  
A:Accession: S67961  
A:Molecule type: protein  
A:Residues: 533-539;542-551;'AQ' 619-634 <NEM>  
C:Comment: In response to temperature stress, to treatment with certain chemicals and am  
molecular-weight proteins.  
C:Genetics:

A:Gene: GDB:HSPCA; HSPC1  
A:Cross-references: GDB:118813; OMIM:140571  
A:Map position: Xpter-Xq22  
A:Introns: 54/3; 177/1; 221/3; 327/3; 383/1; 446/3; 496/1; 585/3; 697/1  
C:Superfamily: heat shock protein 90  
C:Keywords: estrogen-induced protein; heat shock; phosphoprotein; steroid receptor co  
F:223-322/Region: highly charged  
F:534-569/Region: highly charged  
F:231,263/Binding site: phosphate (Ser) (covalent) #status experimental  
Query Match 71.2%; Score 37; DB 1; Length 732;  
Best Local Similarity 80.0%; Pred. No. 79;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KKKKKKFK 10  
HHU86  
Db 274 KKKKKKIK 283  
RESULT 15  
HMS86  
heat shock protein 86 - mouse  
N:Alternate names: HSP86  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 26-Feb-1999  
C:Accession: B32848; PS0021; A40338; A37345  
R:Moore, S.K.; Kozak, C.; Robinson, E.A.; Ullrich, S.J.; Appella, E.  
J. Biol. Chem. 264, 5343-5351, 1989  
A:Title: Murine 86- and 84-kDa heat shock proteins, cDNA sequences, chromosome assign  
A:Reference number: A32848; MUID:89174568  
A:Accession: B32848  
A:Molecule type: mRNA  
A:Residues: 1-733 <MR1>  
A:Cross-references: GB:J04633; NID:g194030; PID:g309318  
R:Hoffmann, T.; Hovemann, B.  
Gene 74, 491-501, 1988  
A:Title: Heat-shock proteins, Hsp84 and Hsp86, of mice and men: two related genes enc  
A:Reference number: PS0020; MUID:89232740  
A:Accession: PS0021  
A:Molecule type: mRNA  
A:Residues: 6,'A',8-242,247-355,'K' <HOF>  
A:Cross-references: GB:M36830; NID:g194032; PID:g194033  
R:Moore, S.K.; Appella, E.; Villar, C.J.; Kozak, C.A.  
Genomics 10, 1019-1029, 1991  
A:Title: Mapping of the mouse 86-kDa heat-shock protein expressed gene (Hsp86-1) on c  
A:Reference number: A40338; MUID:92009901  
A:Accession: A40338  
A:Molecule type: DNA  
A:Residues: 556-634 <MR2>  
A:Cross-references: GB:M57673; NID:g194028; PID:g194029  
R:Legagneux, V.; Mezger, V.; Quelard, C.; Barnier, J.V.; Bensaude, O.; Morange, M.  
Differentiation 41, 42-48, 1989  
A:Title: High constitutive transcription of HSP86 gene in murine embryonal carcinoma  
A:Reference number: A37345; MUID:90033873  
A:Accession: A37345  
A:Molecule type: mRNA  
A:Residues: 450-733 <LEG>  
A:Cross-references: GB:X16857; NID:g51456; PID:g51457  
C:Comment: In response to temperature stress, to treatment with certain chemicals and  
molecular-weight proteins.  
C:Genetics: This protein is one of two forms of 80-90 kDa heat shock proteins found in  
A:Gene: HSP86  
A:Introns: 586/3  
A:Note: the list of introns may be incomplete  
C:Superfamily: heat shock protein 90  
C:Keywords: estrogen-induced protein; heat shock; phosphoprotein; steroid receptor co  
F:2-733/Product: heat shock protein 86 #status experimental <MAT>  
F:223-323/Region: highly charged  
F:535-570/Region: highly charged

Query Match 71.2%; Score 37; DB 1; Length 733;  
Best Local Similarity 80.0%; Pred. No. 79;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KKKKKKFKFK 10  
      ||||| |  
Db 275 KKKKKIKIK 284

Search completed: September 7, 1999, 23:06:28  
Job time: 2476 sec

GenCore version 4.5

Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:35 ; Search time 80.79 Seconds  
 (without alignments)  
 1.221 Million cell updates/sec

Title: US-09-124-280A-12

Perfect score: 52

Sequence: 1 KKKKKKFKF 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/PCTUS9\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	1	US-08-097-830E-12
2	52	100.0	10	2	US-08-456-112B-12
3	44	84.6	332	2	US-08-405-175A-5
4	44	84.6	335	2	US-08-405-175A-6
5	44	84.6	309	2	US-08-405-175A-7
6	44	84.6	309	2	US-08-405-175A-8
7	44	84.6	281	2	US-08-405-175A-9
8	40	76.9	10	2	US-08-456-112B-42
9	37	71.2	732	2	US-08-533-669A-18
10	37	71.2	199	2	US-08-405-175A-3
11	37	71.2	200	2	US-08-405-175A-4
12	37	71.2	63	3	PCT-US94-01202-2
13	36	69.2	18	2	US-08-683-877-2
14	35	67.3	16	1	US-08-231-730A-36
15	35	67.3	21	1	US-08-231-730A-37
16	35	67.3	27	1	US-08-231-730A-38
17	35	67.3	16	1	US-08-427-001C-36
18	35	67.3	21	1	US-08-427-001C-37
19	35	67.3	27	1	US-08-427-001C-38
20	35	67.3	16	1	US-08-457-798-36
21	35	67.3	21	1	US-08-457-798-37
22	35	67.3	27	1	US-08-457-798-38
23	35	67.3	16	2	US-08-457-171-36
24	35	67.3	21	2	US-08-457-171-37
25	35	67.3	27	2	US-08-457-171-38
26	35	67.3	16	3	PCT-US94-06176-36
27	35	67.3	21	3	PCT-US94-06176-37
28	35	67.3	27	3	PCT-US94-06176-38
29	35	67.3	16	3	PCT-US94-12550-36
30	35	67.3	21	3	PCT-US94-12550-37
31	35	67.3	27	3	PCT-US94-12550-38
32	35	67.3	16	3	PCT-US95-04335-36
33	35	67.3	21	3	PCT-US95-04335-37
34	35	67.3	27	3	PCT-US95-04335-38
35	35	67.3	16	3	PCT-US95-04718-36
36	35	67.3	21	3	PCT-US95-04718-37
37	35	67.3	27	3	PCT-US95-04718-38
38	35	67.3	16	3	PCT-US95-09338-36
39	35	67.3	21	3	PCT-US95-09338-37

## ALIGNMENTS

## RESULT 1

US-08-097-830E-12  
 ; Sequence 12, Application US/08097830E  
 ; Patent No. 5652211

; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: Peptides For Neutralizing The  
 ; TITLE OF INVENTION: Toxicity of Lipid A  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036

COMPUTER READABLE FORM: 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-003

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-097-830E-12

Query Match 100.0%; Score 52; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10

Db 1 KKKKKKFKF 10

## RESULT 2

US-08-456-112B-12  
 ; Sequence 12, Application US/08456112B  
 ; Patent No. 5834430

; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan

Sequence 38, Appl  
 Sequence 36, Appl  
 Sequence 37, Appl  
 Sequence 38, Appl  
 Sequence 15, Appl  
 Sequence 20, Appl

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-12

Query Match 100.0%; Score 52; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
Db 1 KKKKKKFKK 10

RESULT 3  
US-08-405-175A-5  
Sequence 5, Application US/08405175A  
Patent No. 5885772  
GENERAL INFORMATION:  
APPLICANT: Aderem, Alan A.  
APPLICANT: Chen, Jianmin  
APPLICANT: Chang, Sandy  
TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,175A  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-121A  
TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 332 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: predicted primary structure of human MARCKS  
HYPOTHETICAL: NO  
US-08-405-175A-5

Query Match 84.6%; Score 44; DB 2; Length 332;  
Best Local Similarity 80.0%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
Db 152 KKKKKRFSFK 161

RESULT 4  
US-08-405-175A-6  
Sequence 6, Application US/08405175A  
Patent No. 5885772  
GENERAL INFORMATION:  
APPLICANT: Aderem, Alan A.  
APPLICANT: Chen, Jianmin  
APPLICANT: Chang, Sandy  
TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,175A  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-121A  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 335 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: predicted primary structure of bovine MARCKS  
HYPOTHETICAL: NO  
US-08-405-175A-6

Query Match 84.6%; Score 44; DB 2; Length 335;  
Best Local Similarity 80.0%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFK 10  
DB 151 KKKKKRFSK 160

RESULT 5  
US-08-405-175A-7  
; Sequence 7, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianmin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,175A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-121A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 309 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: predicted primary structure of murine MARCKS  
; HYPOTHETICAL: NO  
US-08-405-175A-7

Query Match 84.6%; Score 44; DB 2; Length 309;  
Best Local Similarity 80.0%; Pred. No. 7.2;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFK 10  
DB 145 KKKKKRFSK 154

RESULT 6  
US-08-405-175A-8  
; Sequence 8, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianmin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,175A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-121A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 309 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: predicted primary structure of murine MARCKS  
; HYPOTHETICAL: NO  
US-08-405-175A-7

QY 1 KKKKKKFK 10  
DB 145 KKKKKRFSK 154

RESULT 7  
US-08-405-175A-9  
; Sequence 9, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianmin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,175A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-121A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 309 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: predicted primary structure of rat MARCKS  
; HYPOTHETICAL: NO  
US-08-405-175A-8

Query Match 84.6%; Score 44; DB 2; Length 309;  
Best Local Similarity 80.0%; Pred. No. 7.2;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFK 10  
DB 145 KKKKKRFSK 154

RESULT 8  
US-08-405-175A-10  
; Sequence 10, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianmin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,175A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-121A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 309 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: predicted primary structure of human MARCKS  
; HYPOTHETICAL: NO  
US-08-405-175A-10

TELEPHONE: 201 487-5800  
 TELEFAX: 201 343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 281 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 DESCRIPTION: predicted primary structure of chicken MARCKS  
 HYPOTHETICAL: NO  
 US-08-405-175A-9

Query Match 84.6%; Score 44; DB 2; Length 281;  
 Best Local Similarity 80.0%; Pred. No. 6.7;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10  
 |||||: ||  
 Db 117 KKKKKRFSFK 126

RESULT 8  
 US-08-456-112B-42  
 Sequence 42, Application US/08456112B  
 Patent No. 5834430  
 GENERAL INFORMATION:  
 APPLICANT: Porro, Massimo  
 TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hedman, Gibson & Costigan  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: LEADING EDGE 486  
 OPERATING SYSTEM: DOS  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,112B  
 FILING DATE: May 31, 1995  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Costigan, James V.  
 REGISTRATION NUMBER: 25,669  
 REFERENCE/DOCKET NUMBER: 576-004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 302-8989  
 TELEFAX: (212) 302-8998  
 INFORMATION FOR SEQ ID NO: 42:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 10 amino acids  
 TYPE: amino acid  
 TOPOLOGY: circular  
 US-08-456-112B-42

Query Match 76.9%; Score 40; DB 2; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 1.7;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9  
 |||||: ||  
 Db 1 KKKKKRFLF 9

RESULT 9  
 US-08-533-669A-18  
 Sequence 18, Application US/08533669A  
 Patent No. 5834592  
 GENERAL INFORMATION:  
 APPLICANT: Corixa Corporation  
 TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
 THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/533,669A  
 FILING DATE: 22-SEP-1995  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.420  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 732 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-533-669A-18

Query Match 71.2%; Score 37; DB 2; Length 732;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10  
 |||||: ||  
 Db 274 KKKKKRKEK 283

RESULT 10  
 US-08-405-175A-3  
 Sequence 3, Application US/08405175A  
 Patent No. 5885772  
 GENERAL INFORMATION:  
 APPLICANT: Aderem, Alan A.  
 APPLICANT: Chen, Jianmin  
 APPLICANT: Chang, Sandy  
 TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Klauber & Jackson  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/405,175A  
;; FILING DATE: 16-MAR-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 600-1-121A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 199 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; DESCRIPTION: rabbit alveolar macrophage MacMARCKS  
US-08-405-175A-3

Query Match 71.2%; Score 37; DB 2; Length 199;  
Best Local Similarity 87.5%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKKFK 10  
| | | | |  
Db 89 KKKKFSFK 96

RESULT 11  
US-08-405-175A-4  
;; Sequence 4, Application US/08405175A  
;; Patent No. 5885772  
;; GENERAL INFORMATION:  
;; APPLICANT: Agere, Alan A.  
;; APPLICANT: Chen, Jianmin  
;; APPLICANT: Chang, Sandy  
;; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/405,175A  
;; FILING DATE: 16-MAR-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 600-1-121A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 200 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; DESCRIPTION: murine resident peritoneal macrophage MacMARCKS

US-08-405-175A-4

Query Match 71.2%; Score 37; DB 2; Length 200;  
Best Local Similarity 87.5%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKKFK 10  
| | | | |  
Db 89 KKKKFSFK 96

RESULT 12  
PCR-US94-01202-2  
;; Sequence 2, Application PC/TUS9401202  
;; GENERAL INFORMATION:  
;; APPLICANT: Ribeiro, Jose M.  
;; APPLICANT: Lerner, Ethan A.  
;; APPLICANT: Remold, Heinz G.  
;; APPLICANT: Titus, Richard G.  
;; APPLICANT: Tsuji, Yoshiharu  
;; APPLICANT: Hansawa, Chika  
;; APPLICANT: Uzuka, Makoto  
;; TITLE OF INVENTION: Method for Enhancing Hair Growth and  
;; TITLE OF INVENTION: Hair Revitalizing Compositions Comprising Lutzomyia  
;; TITLE OF INVENTION: Protein  
;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
;; ADDRESSEE: Dunner  
;; STREET: 1300 I Street, N.W., Suite 700  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20005-3315  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US94/01202  
;; FILING DATE: 10-FEB-1994  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/017,061  
;; FILING DATE: 12-FEB-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meyers, Kenneth J.  
;; REGISTRATION NUMBER: 25,146  
;; REFERENCE/DOCKET NUMBER: 05136-0002-00000  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-408-4000  
;; TELEFAX: 202-408-4400  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 63 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
PCR-US94-01202-2

Query Match 71.2%; Score 37; DB 3; Length 63;  
Best Local Similarity 87.5%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKFK 8  
| : | | | | |  
Db 53 KKKKKFK 50

RESULT 13

```
US-08-683-877-2
; Sequence 2, Application US/08683877
; Patent No. 5776689
; GENERAL INFORMATION:
; APPLICANT: Karin, Michael
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Aronheim, Ami
; TITLE OF INVENTION: Protein Recruitment System
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; City: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,877
; FILING DATE: 19-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-683-877-2

Query Match 59.2%; Score 36; DB 2; Length 18;
Best Local Similarity 80.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFKFK 10
Db 5 KKKKKSKTK 14

RESULT 14
US-08-231-730A-36
; Sequence 36, Application US/08231730A
; Patent No. 5561107
; GENERAL INFORMATION:
; APPLICANT: JAYNES, JESSE M.
; APPLICANT: JULIAN, GORDON R.
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST AN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; CITY: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0

US-08-683-877-2
; Sequence 2, Application US/08683877
; Patent No. 5776689
; GENERAL INFORMATION:
; APPLICANT: Karin, Michael
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Aronheim, Ami
; TITLE OF INVENTION: Protein Recruitment System
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; City: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,877
; FILING DATE: 19-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-683-877-2

Query Match 67.3%; Score 35; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKKFKK 10
Db 1 KKKKKFKK 8

RESULT 15
US-08-231-730A-37
; Sequence 37, Application US/08231730A
; Patent No. 5561107
; GENERAL INFORMATION:
; APPLICANT: JAYNES, JESSE M.
; APPLICANT: JULIAN, GORDON R.
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; CITY: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,730A
; FILING DATE: 04-20-94
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: HULTQUIST, STEVEN J.
; REGISTRATION NUMBER: 28021
; REFERENCE/DOCKET NUMBER: 4013-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-231-730A-36

Query Match 67.3%; Score 35; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKKFKK 10
Db 1 KKKKKFKK 8

RESULT 15
US-08-231-730A-37
; Sequence 37, Application US/08231730A
; Patent No. 5561107
; GENERAL INFORMATION:
; APPLICANT: JAYNES, JESSE M.
; APPLICANT: JULIAN, GORDON R.
; TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; CITY: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,730A
; FILING DATE: 04-20-94
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
```



; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: HULTQUIST, STEVEN J.
; REGISTRATION NUMBER: 28021
; REFERENCE/DOCKET NUMBER: 4013-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-231-730A-37

Query Match 67.3%; Score 35; DB 1; Length 21;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 KKKKFRFK 10
Db 1 KKKKFRVK 8

Search completed: September 7, 1999, 22:38:35
Job time: 7919 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:13 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-12

Perfect score: 52  
Sequence: 1 KKKKKKFK 10

Scoring table: BLOSUM62

Searched: 189963 seqs, 23686106 residues

Database : A\_Geneseq.36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	52	100.0	1 R71783	Peptide neutralisi
2	52	100.0	1 W21600	Antibiotic potenti
3	44	84.6	330 1 R0528	High density lipop
4	44	84.6	20 1 W45880	Peptide membrane b
5	44	84.6	25 1 W71666	MARCKS-derived tar
6	43	82.7	13 1 W24440	Nucleic acid (NA)
7	42	80.8	24 1 W71667	MacMARCKS-derived
8	40	76.9	10 1 W21630	Antibiotic potenti
9	39	75.0	13 1 R57406	Peptide for treati
10	39	75.0	140 1 R66396	Human psychosis pr
11	39	75.0	9 1 W42478	Protein kinase C p
12	38	73.1	34 1 R11514	Anti-ATLA antibody
13	38	73.1	17 1 W24448	Nucleic acid (NA)
14	38	73.1	45 1 W24450	Nucleic acid (NA)
15	38	73.1	36 1 W24457	Nucleic acid (NA)
16	38	73.1	48 1 W24459	Nucleic acid (NA)
17	38	73.1	24 1 W24443	Nucleic acid (NA)
18	38	73.1	25 1 W24444	Nucleic acid (NA)
19	38	73.1	13 1 W24445	Nucleic acid (NA)
20	38	73.1	14 1 W24446	Nucleic acid (NA)
21	38	73.1	15 1 W24447	Nucleic acid (NA)
22	38	73.1	46 1 W24420	Nucleic acid deliv
23	38	73.1	13 1 W24401	K8, a nucleic acid
24	38	73.1	29 1 W22829	MICFTR based chann
25	38	73.1	30 1 W22830	MICFTR based chann
26	38	73.1	19 1 W38793	Delivery peptide u
27	38	73.1	20 1 W38794	Delivery peptide u
28	38	73.1	21 1 W38795	Delivery peptide u
29	38	73.1	22 1 W38796	Delivery peptide u
30	38	73.1	23 1 W38797	Delivery peptide u
31	38	73.1	24 1 W38798	Delivery peptide u
32	38	73.1	25 1 W38799	Delivery peptide u
33	38	73.1	26 1 W38800	Delivery peptide u
34	38	73.1	27 1 W38801	Delivery peptide u
35	38	73.1	28 1 W38802	Delivery peptide u
36	38	73.1	29 1 W38803	Delivery peptide u
37	38	73.1	30 1 W38804	Delivery peptide u
38	38	73.1	31 1 W38805	Delivery peptide u
39	38	73.1	32 1 W38806	Delivery peptide u
40	38	73.1	33 1 W38807	Delivery peptide u
41	38	73.1	34 1 W38808	Delivery peptide u
42	38	73.1	10 1 W38784	Delivery peptide u
43	38	73.1	11 1 W38785	Delivery peptide u

44	38	73.1	12 1 W38786.	Delivery peptide u					
45	38	73.1	13 1 W38787.	Delivery peptide u					
ALIGNMENTS									
RESULT	1								
ID	R71783								
AC	R71783 standard; peptide; 10 AA.								
DT	01-OCT-1995 (first entry)								
DE	Peptide neutralising toxicity of Lipid A.								
KW	Endotoxin; Lipid A; septic shock; lipopolysaccharide.								
OS	Synthetic.								
PN	WO9503327-A.								
PD	02-FEB-1995.								
PF	21-JUL-1994; E02413.								
PR	26-JUL-1993; US-097830.								
PA	(BIOS-) BIOSYNTH SRL.								
PI	Porro M;								
DR	WPI; 95-075190/10.								
PT	New peptide(s) for neutralising LPS endotoxin - comprising								
PT	repeating units of a basic aminoacid or basic and hydrophobic								
PT	amino acids								
PS	Claim 14; Page 21; 26pp; English.								
CC	New peptides are claimed which are linear or cyclic peptides which								
CC	include units of formula: (A) <sub>n</sub> , where A is the cationic amino acid Lys								
CC	or Arg and n is 7-10; (AB) <sub>m</sub> , where A is Lys or Arg, B is the hydrophobic								
CC	amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and								
CC	(ABC) <sub>p</sub> , where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or								
CC	Trp, and p is 2 or greater.								
CC	The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic								
CC	activity. Hence they can be used therapeutically to treat septic shock								
CC	and also in vitro to detoxify vaccines, drug solutions, injectable								
CC	nutrient solutions, etc.								
CC	The present sequence is a specifically claimed example of the new								
CC	peptides.								
SQ	Sequence 10 AA;								
Query Match	100.0%;	Score 52;	DB 1;	Length 10;					
Best Local Similarity	100.0%;	Pred. No. 0.017;							
Matches	10;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1 KKKKKKFKFK 10								
DB	1 KKKKKKFKFK 10								
RESULT	2								
ID	W21600								
AC	W21600 standard; peptide; 10 AA.								
DT	26-AUG-1997 (first entry)								
DE	Antibiotic potentiating peptide #12.								
KW	Potentiator; antibiotic; microbial infection; lipopolysaccharide;								
KW	permeability; outer bacterial membrane.								
OS	Synthetic.								
PN	WO9638163-A1.								
PD	05-DEC-1996.								
PF	29-MAY-1996; E02313.								
PR	31-MAY-1995; US-456112.								
PA	(BIOS-) BIOSYNTH SRL.								
PI	Porro M, Varra M;								
DR	WPI; 97-034095/03.								
PT	Potentiating activity of antibiotic with peptide contg. cationic								
PT	amino acid sequence - reduces dose of antibiotic required								
PS	Claim 16; Page 25; 37pp; English.								
CC	The sequences given in W21589-633 represent peptides which act to								
CC	potentiate the activity of an antibiotic when they are co-administered								
CC	with the antibiotic. Compositions containing these peptides are used								
CC	to treat or prevent microbial infections. These peptides bind to								

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 52; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.017;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
 |||||  
 Db 1 KKKKKKFKK 10

RESULT 3  
 R05528  
 ID R05528 standard; protein; 330 AA.  
 AC R05528;  
 DT 23-OCT-1990 (first entry)  
 DE High density lipoprotein (HDL) binding protein.  
 KW High density lipoprotein; HDL-binding protein; atherosclerosis;  
 KW hypercholesterolaemia; ds.  
 OS Homo sapiens.  
 PN W09005744-A.  
 PD 31-MAY-1990.  
 PF 17-NOV-1989; 005169.  
 PR 18-NOV-1988; US-273388.  
 PA (UNIW) Univ of Washington, (ZYMO-) Zymogenetics Inc.  
 PI Oram JF, McKnight GL, Hart CE, Curtis DA;  
 DR WPI; 90-193405/25.  
 DR N-PSDB; Q04784.  
 PT New mammalian proteins binding high density lipoprotein sub-class 3 -  
 PT DNA encoding them and derived antibodies, for screening  
 PT potentially therapeutic HDL analogues and for diagnosing risk of  
 PT atherosclerosis.  
 PS Claim 4; Fig 1A-D; 79pp; English.  
 CC The protein product may be used to raise Abs, and the cDNA to  
 CC create probes, both useful in screening for HDL analogues,  
 CC agonists and antagonists, and in identifying abnormalities in the  
 CC HDL binding/receptor pathway. HDL analogues can be used in treating  
 CC hypercholesterolaemia and atherosclerosis  
 SQ Sequence 330 AA;

Query Match 84.6%; Score 44; DB 1; Length 330;  
 Best Local Similarity 80.0%; Pred. No. 5.7;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
 |||||  
 Db 152 KKKKKRFSFK 161

RESULT 4  
 W45880  
 ID W45880 standard; peptide; 20 AA.  
 AC W45880;  
 DT 30-JUN-1998 (first entry)  
 DE Peptide membrane binding element.  
 KW Membrane binding element; thrombotic disease; inflammation;  
 KW complement related disease; soluble peptide.  
 OS Synthetic.  
 PN W09802454-A2.  
 PD 22-JAN-1998.  
 PF 08-JUL-1997; E03715.  
 PR 15-JUL-1996; GB-014871.  
 PA (ADPR-) ADPROTECH PLC.  
 PI Dodd I, Mossakowska DEI, Smith RAG;  
 DR WPI; 98-110524/10.  
 PT Derivatives of soluble poly:peptide(s) bonded to low affinity

PT membrane binding groups - useful for treating complement-related and  
 PT thrombotic diseases, providing improved localisation at cellular  
 PT membranes  
 PS Claim 11; Page 70; 75pp; English.  
 CC The present peptide sequence represents a specifically claimed membrane  
 CC binding element. The invention relates to a soluble derivative (A) of a  
 CC soluble polypeptide (I), which comprises at least 2 heterologous  
 CC membrane-binding elements (MBE) of low membrane affinity covalently  
 CC associated with (I). MBE interact, independently and with thermodynamic  
 CC additivity, with components of cellular or artificial membranes exposed  
 CC to extracellular fluids. (A) are used to treat disorders treatable with  
 CC (I) itself, specifically inflammation or any other complement-related  
 CC disorder (e.g. neurological disease, graft rejection, myocardial  
 CC infarction, sepsis, rheumatoid arthritis and many others; including  
 CC application to indwelling devices) and thrombolytic disease, but also to  
 CC treat allergy, induce weight loss, to treat ischaemia or asthma and as  
 CC immuno-modulators for treating multiple sclerosis. (A) are administered  
 CC orally, topically, by injection or inhalation at 0.01-10 (preferably  
 CC 0.1-10) mg/kg/day.  
 SQ Sequence 20 AA;

Query Match 84.6%; Score 44; DB 1; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 0.46;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10  
 |||||  
 Db 8 KKKKKRFSFK 17

RESULT 5  
 W71666  
 ID W71666 standard; Peptide; 25 AA.  
 AC W71666;  
 DT 11-JAN-1999 (first entry)  
 DE MARCKS-derived target peptide.  
 KW Calmodulin; green fluorescent protein; GFP; cameleon;  
 KW fluorescence resonance energy transfer; FRET; calcium; sensor;  
 KW analysis; assay; myristoylated alaminte-rich C kinase substrate;  
 KW MARCKS.  
 OS Synthetic.  
 PN W09840477-A1.  
 PD 17-SEP-1998.  
 PF 13-MAR-1998; U04978.  
 PR 27-AUG-1997; US-919143.  
 PR 14-MAR-1997; US-818252.  
 PR 14-MAR-1997; US-818253.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Miyawaki A, Tsien RY;  
 DR WPI; 98-520809/44.  
 PT New fluorescent protein sensors for detection of analytes -  
 PT comprises a binding protein moiety having an analyte binding region  
 PT and bound donor and acceptor fluorescent protein moieties  
 PS Disclosure; Page 21; 108pp; English.

QY 1 KKKKKKFKK 10  
 |||||  
 Db 152 KKKKKRFSFK 161

RESULT 4  
 W45880  
 ID W45880 standard; peptide; 20 AA.  
 AC W45880;  
 DT 30-JUN-1998 (first entry)  
 DE Peptide membrane binding element.  
 KW Membrane binding element; thrombotic disease; inflammation;  
 KW complement related disease; soluble peptide.  
 OS Synthetic.  
 PN W09802454-A2.  
 PD 22-JAN-1998.  
 PF 08-JUL-1997; E03715.  
 PR 15-JUL-1996; GB-014871.  
 PA (ADPR-) ADPROTECH PLC.  
 PI Dodd I, Mossakowska DEI, Smith RAG;  
 DR WPI; 98-110524/10.  
 PT Derivatives of soluble poly:peptide(s) bonded to low affinity

Query Match 84.6%; Score 44; DB 1; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 0.56;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 KKKKKKFKK 10
    |||||:| |
DB 1 KKKKKRFSFK 10

RESULT 6
W24440
ID W24440 standard; peptide: 13 AA.
DT 30-SEP-1997 (first entry)
DE Nucleic acid (NA) binding peptide used in NA delivery to cells.
KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;
JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.
OS Synthetic.
PN W09640958-A1.
PD 19-DEC-1996.
PF 23-APR-1996; U05679.
PR 07-JUN-1995; US-484777.
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Smith LC, Sparrow JT, Woo SL;
DR WPI; 97-052345/05.
PT Nucleic acid transporter useful in gene therapy - contains binding
complex associated with surface and nuclear ligands and lysis agent
PS Disclosure: Page 49; 125pp; English.
CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both
condensing and stabilising a NA. The peptides can be conjugated to a
lytic peptide to form a nucleic acid transporter system. The lysis agent
forms an alpha-helical structure. The transporter system is used to
deliver nucleic acid to a cell and for treating humans by gene therapy.
CC By taking advantage of the characteristics of both the lysis agents
and the binding molecules, delivery of the nucleic acid is enhanced.
CC Specific lysis agents are capable of releasing the nucleic acid
into the cellular interior from the endosome. Release is efficient
CC without endosomal/lysosomal degradation. Once released the binding
complexes help target the nucleic acid to the nucleus.
SQ Sequence 13 AA;

Query Match 82.7%; Score 43; DB 1; Length 13;
Best Local Similarity 90.0%; Pred. No. 0.44;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 10
    |||||:| |
DB 4 KKKKKKFKK 13

RESULT 7
W71667
ID W71667 standard; Peptide: 24 AA.
DT 11-JAN-1999 (first entry)
DE MacMARCKS-derived target peptide.
KW Calmodulin; green fluorescent protein; GFP; cameleon;
KW Fluorescence resonance energy transfer; FRET; calcium; sensor;
KW analysis; assay; myristoylated alaminte-rich C Kinase substrate;
KW MacMARCKS.
OS Synthetic.
PN W09840477-A1.
PD 17-SEP-1998.
PF 13-MAR-1998; U04978.
PR 27-AUG-1997; US-919143.
PR 14-MAR-1997; US-818252.
PR 14-MAR-1997; US-818253.
PA (REGC ) UNIV CALIFORNIA.
PI Miyawaki A, Tsien RY;
DR WPI; 98-520809/44.
PT New fluorescent protein sensors for detection of analytes -
comprises a binding protein moiety having an analyte binding region
PT and bound donor and acceptor fluorescent protein moieties
PS Disclosure: Page 21; 108pp; English.
CC This peptide represents a target moiety from myristoylated

CC alaminte-rich C kinase substrate (F52 or MacMARCKS) recognised
by calmodulin. The invention provides fluorescent indicators and
methods for using them to determine the concentration of an
analyte, such as calcium ion, in vitro and in vivo. Fluorescent
indicators include a binding protein moiety (e.g. calmodulin) and
donor and acceptor fluorescent protein moieties, preferably derived
from Aquorea green fluorescent protein (see W71645-48). The
binding protein preferably binds target peptides (see W71649-79) in
addition to the analyte. The target peptide moieties can be
modified to enhance the response of the fluorescent indicator to
the analyte.
SQ Sequence 24 AA;

Query Match 80.8%; Score 42; DB 1; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKKKFKK 10
    |||||:| |
DB 1 KKKKKRFSFK 9

RESULT 8
W21630
ID W21630 standard; peptide: 10 AA.
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #42.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
permeability; outer bacterial membrane.
OS Synthetic.
PN W09638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M;
DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
amino acid sequence - reduces dose of antibiotic required
PS Claim 46; Page 29; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
potentiate the activity of an antibiotic when they are co-administered
with the antibiotic. Compositions containing these peptides are used
to treat or prevent microbial infections. These peptides bind to
lipopolysaccharide on the bacteria so may increase permeability of
the outer bacterial membrane to the antibiotic, allowing a reduction
in the dose of antibiotic required by 10-90% of the normal dose for
in vivo or in vitro application. Any toxic side effects are
correspondingly reduced.
SQ Sequence 10 AA;

Query Match 76.9%; Score 40; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKK 9
    |||||:| |
DB 1 KKKKKRFLF 9

RESULT 9
R57406
ID R57406 standard; Protein; 13 AA.
AC R57406;
DT 21-MAR-1995 (first entry)
DE Peptide for treating diseases related to anti-DNA antibodies.
KW Carrier; absorbing agent; treatment; anti-DNA antibody; Immune complex.
OS Synthetic.
PN J06192290-A.
PD 12-JUL-1994.

```

PF 18-JAN-1993; 006098.  
 PR 30-SEP-1992; JP-261821.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI; 94-260510/32.  
 PT A peptide and an adsorbing agent prepd. by immobilising it on a  
 carrier - useful for treatment of diseases related to anti-DNA  
 PT antibodies and immune complexes  
 CC The sequences given in R57386-413 are peptides which are all covered  
 PS Disclosure; Page 13; 14pp; Japanese.  
 CC by the claimed generic formula:  
 CC H-X-(A-B)n-Y-Z  
 CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
 CC B = Trp, Phe, Asn or Glu;  
 CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
 CC consisting of 2-10 residues, provided that at least one of  
 CC X or Y are present;  
 CC Z = OH or NH2; and  
 CC n = 2-5.  
 CC These peptides may be immobilised on a carrier in the preparation of an  
 CC absorbing agent which may be used in the treatment of diseases related  
 CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 13 AA;

Query Match 75.0%; Score 39; DB 1; Length 13;  
 Best Local Similarity 77.8%; Pred. No. 1.7;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKKKKKFK 10  
 DB 1 KKKKFEFE 9  
 RESULT 10  
 R66396  
 ID R66396 standard; Peptide; 140 AA.  
 AC R66396;  
 DT 11-AUG-1995 (first entry)  
 DE Human psychosis protecting peptide encoded by reading frame 3.  
 KW Psychosis protecting peptide.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT misc\_difference 1..140  
 FT /note= "X corresp. to translated stop codon"

PN W09426107-A.  
 PD 24-NOV-1994.  
 PF 13-MAY-1994; U05445.  
 PR 13-MAY-1993; US-060560.  
 PA (UYN ) UNIV NEW YORK STATE.  
 PI Basham DA, Friedhoff AJ, Miller JC;  
 DR WPI; 95-006234/01.  
 DR N-PSDB; Q75090.  
 PT New nucleic acids encoding psychosis protecting peptide and  
 PT antibodies - for the treatment, diagnosis and research of  
 PT psychotic disorders, such as schizophrenia  
 PS Claim 1; Page 57-58; 87 pp; English.  
 CC Psychosis protecting peptide (PP peptide) can be obt'd. initially  
 CC by using the sequence in Q75090 as a basis for designing  
 CC polynucleotide probes to clone, sequence and express or synthesize  
 CC PP related protein and peptides occurring in normal individuals,  
 CC and to a substantially lesser degree in individuals with  
 CC psychotic disorders. The PP encoding gene was discovered using  
 CC the subtraction cloning of cDNA from mRNA obt'd. from monozygotic  
 CC twins discordant for schizophrenia and assaying for clones  
 CC in which expression is greatest in the 'well' twin. A nt  
 CC sequence substantially corresp. to Q75090 is claimed, which  
 CC comprises 30-X nts, where X=80, 95, 138, 222, 243, 249, 260, 295,  
 CC 407 or 423. A PP peptide of 10-141 AAs is also claimed.  
 SQ Sequence 140 AA;

Query Match 75.0%; Score 39; DB 1; Length 140;  
 Best Local Similarity 88.9%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKKKKKFK 9  
 DB 5 KKKKKKFK 13  
 RESULT 11  
 W42478  
 ID W42478 standard; Protein; 9 AA.  
 AC W42478;  
 DT 15-SEP-1998 (first entry)  
 DE Protein kinase C phosphorylation site motif #1.  
 KW Green fluorescent protein; protein kinase; phosphate donor;  
 KW phosphorylation; drug screening; receptor-ligand binding; signalling;  
 KW protein-protein interaction; kinase activation.  
 OS Unknown.  
 PN W09802571-A1.  
 PD 22-JAN-1998.  
 PF 16-JUL-1997; U12410.  
 PR 16-JUL-1996; US-680877.  
 PR 16-JUL-1996; US-679865.  
 PR 16-JUL-1996; US-680876.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Cubitt AB, Tsien RY;  
 DR WPI; 98-110616/10.  
 PT Assays for protein kinase and modulators - using a fluorescent  
 PT protein substrate which exhibits a different fluorescent property in  
 PT the phosphorylated and un-phosphorylated state  
 PS Claim 26; Page 19; 65pp; English.  
 CC This sequence represents a protein kinase C phosphorylation site motif  
 CC which is used to design a method which determines if a sample contains  
 CC protein kinase (PK) activity. The method involves contacting the sample  
 CC with a phosphate donor (PD) and a fluorescent protein (FP) substrate for  
 CC a PK, the protein substrate comprising a FP moiety and a phosphorylation  
 CC site for a PK, where the protein substrate exhibits a different  
 CC fluorescent property in the phosphorylated state than in the  
 CC un-phosphorylated state. The protein substrate is then excited and the  
 CC amount of a fluorescent property that differs in the un-phosphorylated  
 CC state and phosphorylated state is measured, whereby an amount that is  
 CC consistent with the presence of the protein substrate in its  
 CC phosphorylated state indicates the presence of PK activity. The method  
 CC and products can be used in drug screening. They can be used for  
 CC screening for compounds which affect cellular events, including  
 CC receptor-ligand binding, protein-protein interactions or kinase  
 CC activation, which signal to the target kinase.  
 SQ Sequence 9 AA;

Query Match 75.0%; Score 39; DB 1; Length 9;  
 Best Local Similarity 77.8%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKKKFK 10  
 DB 1 KKKKFSFK 9  
 RESULT 12  
 R11514  
 ID R11514 standard; Protein; 34 AA.  
 AC R11514;  
 DT 21-JUN-1991 (first entry)  
 DE Anti-ATLA antibody-binding peptide #3  
 KW adult T cell leukaemia associated antigen; HTLV-1; ATL.  
 OS Synthetic.  
 PN EP-423649-A.  
 PD 24-APR-1991.  
 PF 12-OCT-1990; 119624.  
 PR 13-OCT-1989; JP-266983.  
 PA (KURS ) KURARAY KK.  
 PI Maeda Y, Shiraki H, Washitani Y, Kuroda N, Yamada K;  
 PI Oka K, Namba T;

DR WPI: 91-118942/17  
PT New peptide(s) able to bind to anti-ATLA antibody - used in  
PT diagnosis and to treat HTLV-1-associated myelopathy and diseases  
PT caused by ATL  
PS Claim 4: Page 31; 44 pp; English.  
CC This peptide is able to bind to an antibody specific against an  
CC adult T-cell leukaemia associated antigen. The peptide can have  
CC from 1 to 10 lys residues at the N-terminus. The invention also  
CC covers an adsorbent for anti-ATLA antibodies, comprising the peptide  
CC immobilised on a carrier and a reagent containing the peptide, for  
CC measuring the antibody. The reagent and the adsorbent can be used  
CC to treat an adult T-cell leukaemia virus-infectious disease.  
CC See also R11512 and R11513.  
CC Sequence 34 AA;  
SQ

Query Match 73.1%; Score 38; DB 1; Length 34;  
Best Local Similarity 88.9%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 9  
| | | | | | |  
DB 3 KKKKKKKKF 11

RESULT 13  
W24448  
ID W24448 standard; peptide: 17 AA.  
AC W24448;  
DT 30-SEP-1997 (first entry)  
DE Nucleic acid (NA) binding peptide used in NA delivery to cells.  
KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;  
KW JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.  
OS Synthetic.  
PN W09640958-A1.  
PD 19-DEC-1996.  
PF 23-APR-1996; U05679.  
PR 07-JUN-1995; US-484777.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Smith LC, Sparrow JT, Woo SL;  
DR WPI: 97-052345/05.  
PT Nucleic acid transporter useful in gene therapy - contains binding  
PT complex associated with surface and nuclear ligands and lysis agent  
PS Disclosure; Page 49; 125pp; English.  
CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both  
CC condensing and stabilising a NA. The peptides can be conjugated to a  
CC lytic peptide to form a nucleic acid transporter system. The lysis agent  
CC forms an alpha-helical structure. The transporter system is used to  
CC deliver nucleic acid to a cell and for treating humans by gene therapy.  
CC By taking advantage of the characteristics of both the lysis agents  
CC and the binding molecules, delivery of the nucleic acid is enhanced.  
CC Specific lysis agents are capable of releasing the nucleic acid  
CC into the cellular interior from the endosome. Release is efficient  
CC without endosomal/lysosomal degradation. Once released the binding  
CC complexes help target the nucleic acid to the nucleus.  
CC Sequence 17 AA;  
SQ

Query Match 73.1%; Score 38; DB 1; Length 17;  
Best Local Similarity 80.0%; Pred. No. 3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10  
| | | | | | |  
DB 8 KKKKKKKKK 17

RESULT 14  
W24450  
ID W24450 standard; peptide: 45 AA.  
AC W24450;  
DT 30-SEP-1997 (first entry)  
DE Nucleic acid (NA) binding peptide used in NA delivery to cells.

KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;  
KW JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.  
OS Synthetic.  
PN W09640958-A1.  
PD 19-DEC-1996.  
PF 23-APR-1996; U05679.  
PR 07-JUN-1995; US-484777.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Smith LC, Sparrow JT, Woo SL;  
DR WPI: 97-052345/05.  
PT Nucleic acid transporter useful in gene therapy - contains binding  
PT complex associated with surface and nuclear ligands and lysis agent  
PS Disclosure; Page 49; 125pp; English.  
CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both  
CC condensing and stabilising a NA. The peptides can be conjugated to a  
CC lytic peptide to form a nucleic acid transporter system. The lysis agent  
CC forms an alpha-helical structure. The transporter system is used to  
CC deliver nucleic acid to a cell and for treating humans by gene therapy.  
CC By taking advantage of the characteristics of both the lysis agents  
CC and the binding molecules, delivery of the nucleic acid is enhanced.  
CC Specific lysis agents are capable of releasing the nucleic acid  
CC into the cellular interior from the endosome. Release is efficient  
CC without endosomal/lysosomal degradation. Once released the binding  
CC complexes help target the nucleic acid to the nucleus.  
CC Sequence 45 AA;  
SQ

Query Match 73.1%; Score 38; DB 1; Length 45;  
Best Local Similarity 80.0%; Pred. No. 7.2;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFKF 10  
| | | | | | |  
DB 36 KKKKKKKKK 45

RESULT 15  
W24457  
ID W24457 standard; peptide: 36 AA.  
AC W24457;  
DT 30-SEP-1997 (first entry)  
DE Nucleic acid (NA) binding peptide used in NA delivery to cells.  
KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;  
KW JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.  
OS Synthetic.  
PN W09640958-A1.  
PD 19-DEC-1996.  
PF 23-APR-1996; U05679.  
PR 07-JUN-1995; US-484777.  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
PI Smith LC, Sparrow JT, Woo SL;  
DR WPI: 97-052345/05.  
PT Nucleic acid transporter useful in gene therapy - contains binding  
PT complex associated with surface and nuclear ligands and lysis agent  
PS Disclosure; Page 50; 125pp; English.  
CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both  
CC condensing and stabilising a NA. The peptides can be conjugated to a  
CC lytic peptide to form a nucleic acid transporter system. The lysis agent  
CC forms an alpha-helical structure. The transporter system is used to  
CC deliver nucleic acid to a cell and for treating humans by gene therapy.  
CC By taking advantage of the characteristics of both the lysis agents  
CC and the binding molecules, delivery of the nucleic acid is enhanced.  
CC Specific lysis agents are capable of releasing the nucleic acid  
CC into the cellular interior from the endosome. Release is efficient  
CC without endosomal/lysosomal degradation. Once released the binding  
CC complexes help target the nucleic acid to the nucleus.  
CC Sequence 36 AA;  
SQ

Query Match 73.1%; Score 38; DB 1; Length 36;  
Best Local Similarity 80.0%; Pred. No. 5.9;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKKK 10  
| | | | | : |  
Db 27 KKKKKKKK 36

Search completed: September 7, 1999, 20:37:14  
Job time: 18463 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:43 ; Search time 148.39 Seconds  
(without alignments)  
4.562 Million cell updates/sec

Title: US-09-124-280A-13

Perfect score: 79

Sequence: 1 CKCKCKCKCK 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_ordanelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	64	81.0	257	12 Q68385	Q68385 human cytom
2	53.5	67.7	1698	5 Q94438	Q94438 chironomus
3	52.5	66.5	1704	5 Q94446	Q94446 chironomus
4	52	65.8	69	10 Q22433	Q22433 arabadopsis
5	49	62.0	248	5 Q96282	Q96282 plasmodium
6	49	62.0	618	10 Q64758	Q64758 arabadopsis
7	49	62.0	413	12 Q69566	Q69566 human herpe
8	47	59.5	174	5 Q15663	Q15663 caenorhabdi
9	47	59.5	601	5 Q22913	Q22913 caenorhabdi
10	46.5	58.9	73	5 Q62554	Q62554 mytilus edu
11	46.5	58.9	73	5 Q62555	Q62555 mytilus edu
12	46.5	58.9	61	6 Q18842	Q18842 balaena mys
13	46.5	58.9	61	6 P79379	P79379 sus scrofa
14	46.5	58.2	134	5 Q15649	Q15649 plasmodium
15	46	58.2	215	5 Q21162	Q21162 caenorhabdi
16	46	58.2	455	10 Q92VA2	Q92VA2 arabadopsis
17	46	58.2	455	10 Q92VA1	Q92VA1 arabadopsis
18	45	57.0	1131	5 Q18165	Q18165 caenorhabdi
19	45	57.0	676	5 Q44188	Q44188 caenorhabdi
20	45	57.0	80	10 Q22488	Q22488 oryza sativ
21	45	57.0	59	10 Q65192	Q65192 fagopyrum e
22	45	57.0	80	10 Q44107	Q44107 oryza sativ
23	45	57.0	285	11 Q60501	Q60501 cricetus
24	44	55.7	322	5 P90550	P90550 leishmania
25	44	55.7	143	5 Q15660	Q15660 plasmodium
26	44	55.7	75	5 Q96388	Q96388 perna virid
27	44	55.7	815	10 Q81833	Q81833 arabadopsis
28	43.5	55.1	254	4 Q16889	Q16889 homo sapien
29	43.5	55.1	191	4 Q75875	Q75875 homo sapien

```

30 43.5 55.1 209 4 Q60720 060720 homo sapien
31 43.5 55.1 48 6 P79380 079380 sus scrofa
32 43.5 55.1 190 6 Q77643 077643 ovis aries
33 43.5 55.1 102 11 Q63672 063672 rattus norv
34 43.5 55.1 110 11 Q88911 088911 rattus norv
35 43.5 55.1 188 13 Q73682 073682 brachydanio
36 43.5 55.1 60 13 Q42152 042152 anbystoma m
37 43.5 55.1 194 13 Q42572 042572 xenopus lae
38 43 54.4 2796 4 Q95071 095071 homo sapien
39 43 54.4 188 5 Q18238 018238 caenorhabdi
40 43 54.4 73 5 Q76953 076953 lumbricus c
41 43 54.4 79 5 Q76956 076956 lumbricus r
42 43 54.4 164 5 Q22048 022048 caenorhabdi
43 43 54.4 74 5 Q96822 096822 eisenia foe
44 43 54.4 625 10 Q22925 022925 arabidopsis
45 43 54.4 63 10 Q49125 049125 fritillaria

```

#### ALIGNMENTS

```

RESULT 1
Q68385 PRELIMINARY; PRT; 257 AA.
AC Q68385;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ORF UL133.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TOLEDO;
RX MEDLINE; 96099416.
RA CHA T.A., TOM E., KEMBLE G.W., DUKE G.M., MOCARSKI E.S., SPATTE R.R.;
RT "Human cytomegalovirus clinical isolates carry at least 19 genes not
RT found in laboratory strains.";
RL J. Virol. 70:78-83(1996).
DR EMBL: U33331; AAA85872.1; -.
SQ SEQUENCE 257 AA; 27470 MW; EC259DCB CRC32;

```

```

Query Match 81.0%; Score 64; DB 12; Length 257;
Best Local Similarity 72.7%; Pred. No. 0.0013;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKCKCKCKCK 11
Db 87 CSCCKCKCKC 97

RESULT 2
Q94438 PRELIMINARY; PRT; 1698 AA.
ID Q94438;
AC Q94438;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE 185 KDA SILK PROTEIN.
OS Chironomus pallidivittatus (Midge).
OS Chironomidae; Chirodoma.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Diptera; Nematocera; Chironomoidea; Chironomidae;
OC Chironominae; Chironomus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLAND;
RA CASE S.T., COX C., BELL W.C., HOFFMAN R.T., MARTIN J., HAMILTON R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U54640; AAA99803.1; -.
SQ SEQUENCE 1698 AA; 186164 MW; 858E520 CRC32;

```

Query Match 67.7%; Score 53.5; DB 5; Length 1698;  
Best Local Similarity 34.6%; Pred. No. 0.16;  
Matches 9; Conservative 2; Mismatches 15; Gaps 1;

QY 1 CKCKCK-----CKCKC 11  
|:||||| |:|||  
Db 589 CECKCKNNPCTSPQVWDADDCCKC 614

RESULT 3  
Q94446 PRELIMINARY; PRT; 1704 AA.  
AC Q94446;

DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE 220 KDA SILK PROTEIN.  
GN SP220.  
OS Chironomus thummi thummi (Midge).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Diptera; Nematocera; Chironomidae; Chironomidae;  
OC Chironominae; Chironomus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-SALIVARY GLAND;  
RA CASE S.T., COX C., BELL W.C., HOFFMAN R.T., MARTIN J., HAMILTON R.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U54641; AAA99804.1; -.  
SQ SEQUENCE 1704 AA; 185745 MW; 0245A38E CRC32;

Query Match 66.5%; Score 52.5; DB 5; Length 1704;  
Best Local Similarity 34.6%; Pred. No. 0.23;  
Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 1;

QY 1 CKCKCK-----CKCKC 11  
|:||||| |:|||  
Db 594 CECKCKSNPTCSDPOIWDADDCCKC 619

RESULT 4  
O22433 PRELIMINARY; PRT; 69 AA.  
AC O22433;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE METALLOTHIONEIN-LIKE PROTEIN.  
GN MT3.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA BUNDITHA W., GOLDSBROUGH P.B.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF013959; AAB67234.1; -.  
DR PFAM; PF01439; Metallothio\_2; 1.  
SQ SEQUENCE 69 AA; 7373 MW; CE5C6598 CRC32;

Query Match 65.8%; Score 52; DB 10; Length 69;  
Best Local Similarity 63.6%; Pred. No. 0.038;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
|:||||| |:|  
Db 51 CKCKGSSCSC 61

RESULT 5  
O95282 PRELIMINARY; PRT; 248 AA.  
AC O95282;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE PREDICTED INTEGRAL MEMBRANE PROTEIN.  
GN PFB0950W.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
falciparum".  
RL Science 282:1126-1132(1998).  
DR EMBL; AE001428; AAC71979.1; -.  
SQ SEQUENCE 248 AA; 28259 MW; 4F508144 CRC32;

Query Match 62.0%; Score 49; DB 5; Length 248;  
Best Local Similarity 54.5%; Pred. No. 0.24;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
|:||||| |:|||  
Db 173 CSQLCSCICSC 183

RESULT 6  
O64758 PRELIMINARY; PRT; 618 AA.  
ID O64758;  
AC O64758;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE PUTATIVE VACUOLAR SORTING RECEPTOR.  
GN F1913.17.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,  
RA SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,  
RA SOMERVILLE C.R., VENTER J.C.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC004238; AAC12834.1; -.  
DR PROSITE; PS01187; EGF\_CA; 1.  
KW Glycoprotein; EGF-like domain.  
SQ SEQUENCE 618 AA; 69256 MW; 5759ACA5 CRC32;

Query Match 62.0%; Score 49; DB 10; Length 618;  
Best Local Similarity 52.6%; Pred. No. 0.41;  
Matches 10; Conservative 0; Mismatches 1; Indels 8; Gaps 2;

QY 1 CKCK-CKCK-----CKC 11  
|:||||| |:|||  
Db 522 CKCDGCKCKNNWGGYECKC 540

RESULT 7

Q69566 PRELIMINARY; PRT; 413 AA.  
 AC Q69566;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DE 01-NOV-1998 (TRENBLREL. 08, Last annotation update)  
 DE (HHV-6) U1102, VARIANT A DNA, COMPLETE VIRION GENOME.  
 GN U88.  
 OS Human herpesvirus 6.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Roseolovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93266321.  
 RA GOMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,  
 RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.;  
 RT "The DNA sequence of human herpesvirus-6: structure, coding content,  
 RT and genome evolution.";  
 RL Virology 209:29-51(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 90080132.  
 RA LAWRENCE G.L., CHEE M., CRAXTON M.A., GOMPELS U.A., HONESS R.W.,  
 RA BARRELL B.G.;  
 RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";  
 RL J. Virol. 64:287-299(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 91237802.  
 RA CHANG C.K., BALACHANDRAN N.;  
 RT "Identification, characterization, and sequence analysis of a cDNA  
 RT encoding a phosphoprotein of human herpesvirus 6.";  
 RL J. Virol. 65:2884-2894(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 91333007.  
 RA TEO I.A., GRIFFIN B.E., JONES M.D.;  
 RT "Characterization of the DNA polymerase gene of human herpesvirus  
 RT 6.";  
 RL J. Virol. 65:4670-4680(1991).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 91226542.  
 RA THOMSON B.J., EFSTATHIOU S., HONESS R.W.;  
 RT "Acquisition of the human adeno-associated virus type-2 rep gene by  
 RT human herpesvirus type-6.";  
 RL Nature 351:78-80(1991).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 91374590.  
 RA MARTIN M.E.D., NICHOLAS J., THOMSON B.J., NEWMAN C., HONESS R.W.;  
 RT "Identification of a transactivating function mapping to the putative  
 RT immediate-early locus of human herpesvirus 6.";  
 RL J. Virol. 65:5381-5390(1991).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 92333249.  
 RA EFSTATHIOU S., LAWRENCE G.L., BROWN C.M., BARRELL B.G.;  
 RT "Identification of homologues to the human cytomegalovirus US22 gene  
 RT family in human herpesvirus 6.";  
 RL J. Gen. Virol. 73:1661-1671(1992).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 92148942.  
 RA GENG Y., CHANDRAN B., JOSEPHS S.F., WOOD C.;

RT "Identification and characterization of a human herpesvirus 6 gene  
 RT segment that trans activates the human immunodeficiency virus type 1  
 RT promoter.";  
 RL J. Virol. 66:1564-1570(1992).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93091236.  
 RA GOMPELS U.A., CARSS A.L., SUN N., ARRARD J.R.;  
 RT "Infectivity determinants encoded in a conserved gene block of human  
 RT herpesvirus-6.";  
 RL DNA Seq. 3:25-39(1992).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 92260671.  
 RA NEIPEL F., ELLINGER K., FLECKENSTEIN B.;  
 RT "Gene for the major antigenic structural protein (p100) of human  
 RT herpesvirus 6.";  
 RL J. Virol. 66:3918-3924(1992).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 92333248.  
 RA THOMSON B.J., HONESS R.W.;  
 RT "The right end of the unique region of the genome of human  
 RT herpesvirus 6 U1102 contains a candidate immediate early gene  
 RT enhancer and a homologue of the human cytomegalovirus US22 gene  
 RT family.";  
 RL J. Gen. Virol. 73:1649-1660(1992).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93187613.  
 RA ELLINGER K., NEIPEL F., FOA-TOMASI L., CAMPADELLI-FIUME G.,  
 RA FLECKENSTEIN B.;  
 RT "The glycoprotein B homologue of human herpesvirus 6.";  
 RL J. Gen. Virol. 74:495-500(1993).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93224882.  
 RA GOMPELS U.A., CARRIGAN D.R., CARSS A.L., ARNO J.;  
 RT "Two groups of human herpesvirus 6 identified by sequence analyses of  
 RT laboratory strains and variants from Hodgkin's lymphoma and bone  
 RT marrow transplant patients.";  
 RL J. Gen. Virol. 74:613-622(1993).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93389439.  
 RA LIU D.X., GOMPELS U.A., NICHOLAS J., LELLIOTT C.;  
 RT "Identification and expression of the human herpesvirus 6  
 RT glycoprotein H and interaction with an accessory 40K glycoprotein.";  
 RL J. Gen. Virol. 74:1847-1857(1993).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 94025558.  
 RA LIU D.X., GOMPELS U.A., FOA-TOMASI L., CAMPADELLI-FIUME G.;  
 RT "Human herpesvirus-6 glycoprotein H and L homologs are components of  
 RT the gp100 complex and the gH external domain is the target for  
 RT neutralizing monoclonal antibodies.";  
 RL Virology 197:12-22(1993).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-U1102, VARIANT A;  
 RX MEDLINE; 93331710.  
 RA PELLET P.E., SANCHEZ-MARTINEZ D., DOMINGUEZ G., BLACK J.B., ANTON E.,  
 RA GREENMAYER C., DAMEBAUGH T.R.;  
 RT "A strongly immunoreactive virion protein of human herpesvirus 6  
 RT variant B strain 229: Identification and characterization of the gene  
 RT and mapping of a variant-specific monoclonal antibody reactive



```

QY 1 CKCKCKCKCK 10
   | | | | |
Db 85 CFYCCCKCK 94

RESULT 10
O62554 PRELIMINARY; PRT; 73 AA.
AC O62554;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE METALLOTHIONEIN 10 IB.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytilidae; Mytilus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-DIGESTIVE GLAND;
RA BARSYTE D., WHITE K.N., LOVEJOY D.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ005452; CAA06549.1; -.
DR PFAM; PF00131; metalthio; 1.
SQ SEQUENCE 73 AA; 7220 MW; D437E9F8 CRC32;

Query Match 58.9%; Score 46.5; DB 5; Length 73;
Best Local Similarity 66.7%; Pred. No. 0.27;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCK-CKCKCK 11
   | | | | |
Db 37 CKGSGCKVCKC 48

RESULT 11
O62555 PRELIMINARY; PRT; 73 AA.
AC O62555;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE METALLOTHIONEIN 10 II.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytilidae; Mytilus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-DIGESTIVE GLAND;
RA BARSYTE D., WHITE K.N., LOVEJOY D.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ005453; CAA06550.1; -.
DR PFAM; PF00131; metalthio; 1.
SQ SEQUENCE 73 AA; 7153 MW; 7727E03A CRC32;

Query Match 58.9%; Score 46.5; DB 5; Length 73;
Best Local Similarity 66.7%; Pred. No. 0.27;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCK-CKCKCK 11
   | | | | |
Db 37 CKGSGCKVCKC 48

RESULT 12
O18842 PRELIMINARY; PRT; 61 AA.
AC O18842;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)

```

```

DE METALLOTHIONEIN (MT) (FRAGMENT).
OS Balaena mysticetus (Bowhead whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Cetacea; Mysticeti; Balaenidae; Balaena.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA KAYSER J., O'HARA T., GOODWIN T., LINNEHAN R., HAMMOND T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
CC GLUCOCORTICOIDS.
DE EMBL; AF022117; AAB72006.1; -.
DR PFAM; PF00131; metalthio; 1.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc.
FT METAL 17 17 CLUSTER.
FT METAL 19 19 CLUSTER.
FT METAL 22 22 CLUSTER.
FT METAL 24 24 CLUSTER.
FT METAL 27 27 CLUSTER.
FT METAL 31 31 CLUSTER.
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6025 MW; 1AE6C4FD CRC32;

Query Match 58.9%; Score 46.5; DB 6; Length 61;
Best Local Similarity 64.3%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--CKCK-CKC 11
   | | | | |
Db 13 CTCAGCKCKCKC 26

RESULT 13
P79379 PRELIMINARY; PRT; 61 AA.
AC P79379;
DT 01-MAY-1997 (T-EMBLrel. 03, Created)
DT 01-MAY-1997 (T-EMBLrel. 03, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE METALLOTHIONEIN (MT) (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 98241500.
RA HUANG M.C., PAN P.K., ZENG T.-F., CHEN N.C., PENG J.Y., HUANG P.C.;
RT "Multiple isoforms of metallothionein are expressed in the porcine
RT liver."
RL Gene 211:49-55(1998).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
CC GLUCOCORTICOIDS.
DE EMBL; AB000794; BAA19183.1; -.
DR PFAM; PF00131; metalthio; 1.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc.
FT NON_TER 61 61
FT METAL 24 24 CLUSTER.
FT METAL 27 27 CLUSTER.
FT METAL 17 17 CLUSTER.
FT METAL 19 19 CLUSTER.
FT METAL 22 22 CLUSTER.
FT METAL 31 31 CLUSTER.
SQ SEQUENCE 61 AA; 5970 MW; 43E53AB9 CRC32;

```

Query Match 58.9%; Score 46.5; DB 6; Length 61;  
 Best Local Similarity 64.3%; Pred. No. 0.24;  
 Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
 | | | | | | | | | |  
 DB 13 CTCAGSCKCKDKCK 26

## RESULT 14

O15649 PRELIMINARY; PRT; 134 AA.  
 AC O15649  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
 DE PFEMP1 (FRAGMENT).  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98008145.  
 RA BARUCH D.I., MA X.C., SINGH H.B., BI X., PASLOSKE B.L., HOWARD R.J.;  
 RT "Identification of a region of PfEMP1 that mediates adherence of  
 RT Plasmodium falciparum infected erythrocytes to CD36: conserved  
 RT function with variant sequence.";  
 RL Blood 90:3766-3775(1997).  
 DR EMBL; AF008978; AAB87385.1; .  
 FT NON\_TER 1  
 FT NON\_TER 134 134  
 FT NON\_TER 134 134  
 SQ SEQUENCE 134 AA; 15772 MW; 470CAD51 CRC32;

Query Match 58.2%; Score 46; DB 5; Length 134;  
 Best Local Similarity 63.6%; Pred. No. 0.46;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
 | | | | | | | | | |  
 DB 31 CMSKCNCKCKC 41

## RESULT 15

Q21162 PRELIMINARY; PRT; 215 AA.  
 AC Q21162  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
 DE SIMILAR TO THE DNA-J DOMAIN FOUND IN DNAJ PROTEINS.  
 GN K02G10.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematozoa; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DUREIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;  
 RA LEIMBACH D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U40415; AAB52639.1; .  
 DR PFAM; PF00226; DnaJ; 1.  
 SQ SEQUENCE 215 AA; 24233 MW; 6067B0B9 CRC32;

Query Match 58.2%; Score 46; DB 5; Length 215;  
 Best Local Similarity 54.5%; Pred. No. 0.61;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
 | | | | | | | | | |  
 DB 146 CCGCGMCCQC 156

Search completed: September 7, 1999, 20:34:45  
 Job time: 19741 sec

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	53.5	67.7	1700	1	BAR3_CHITE	Q03376 chironomus
2	51.5	65.2	68	1	MT3_MOUSE	P28184 mus musculus
3	51.5	65.2	66	1	MT3_RAT	P37361 rattus norv
4	51	64.6	256	1	ERG2_NEUCR	Q92254 neurospora
5	50.5	63.9	74	1	MT1_CAEEL	P17511 caenorhabdi
6	50	63.3	71	1	MT21_MYTED	P80251 mytilus edu
7	49	62.0	440	1	G3PT_MOUSE	Q64467 mus musculu
8	48.5	61.4	61	1	MT1B_HUMAN	P07438 homo sapien
9	48.5	61.4	58	1	MT2_CALSI	P55950 callinectes
10	47.5	60.1	61	1	MT1A_RABIT	P11957 oryctolagus
11	47.5	60.1	61	1	MT1G_HUMAN	P13640 homo sapien
12	47.5	60.1	62	1	MT1K_HUMAN	P80296 homo sapien
13	47.5	60.1	61	1	MT1R_HUMAN	Q93083 homo sapien
14	47.5	60.1	61	1	MT1B_RABIT	P80289 oryctolagus
15	47.5	60.1	61	1	MT2D_RABIT	P80291 oryctolagus
16	47.5	60.1	61	1	MT2E_RABIT	P80292 oryctolagus
17	47.5	60.1	61	1	MT2_CRIGR	P02799 cricetulus
18	47.5	60.1	61	1	MT2_MESAU	P17808 mesocricetu
19	47.5	60.1	61	1	MT2_RAT	P04355 rattus norv
20	47.5	60.1	58	1	MT_CARNA	P55948 carlinus ma
21	47.5	60.1	74	1	MT_CRAVI	P23038 crassostrea
22	47.5	60.1	58	1	MT_POIPO	P55952 potamon pot
23	47	59.5	407	1	IE68_HSVSA	Q01042 herpesvirus
24	47	59.5	435	1	YMH1_CAEEL	P34468 caenorhabdi
25	46.5	58.9	72	1	MT11_MYTED	P80246 mytilus edu
26	46.5	58.9	72	1	MT12_MYTED	P80247 mytilus edu
27	46.5	58.9	72	1	MT13_MYTED	P80248 mytilus edu
28	46.5	58.9	72	1	MT14_MYTED	P80249 mytilus edu
29	46.5	58.9	61	1	MT1A_HUMAN	P04731 homo sapien
30	46.5	58.9	61	1	MT1E_HUMAN	P04732 homo sapien
31	46.5	58.9	61	1	MT1F_HUMAN	P04733 homo sapien
32	46.5	58.9	61	1	MT1H_BOVIN	P55942 bos taurus
33	46.5	58.9	61	1	MT1H_HUMAN	P80294 homo sapien
34	46.5	58.9	61	1	MT1I_HUMAN	P80295 homo sapien
35	46.5	58.9	61	1	MT1L_HUMAN	P80297 homo sapien
36	46.5	58.9	61	1	MT1_CANFA	O19000 canis famil
37	46.5	58.9	61	1	MT1_CERAE	P02797 cercopithec
38	46.5	58.9	63	1	MT1_COLL1	P15786 columba liv
39	46.5	58.9	61	1	MT2_BOVIN	P09579 bos taurus
40	46.5	58.9	61	1	MT2_CERAE	P02796 cercopithec
41	46.5	58.9	61	1	MT2_HUMAN	P02795 homo sapien
42	46.5	58.9	61	1	MT2_MOUSE	P02798 mus musculu
43	46.5	58.9	61	1	MT2_STSCO	P14425 stenella co

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 [1]  
 RN TISSUE=RAIN;  
 RC MEDLINE: 92335292.  
 RX PALMITER R.D., FINDLEY S.D., WHITMORE T.E., DURHAM D.M.;  
 RA "MT-III, a brain-specific member of the metallothionein gene family.";  
 RT PROC. NATL. ACAD. SCI. U.S.A. 89:6333-6337(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94314230.  
 RA NARUSE S., IGARASHI S., FURUYA T., KOBAYASHI H., MIYATAKE T.,  
 RA TSUJI S.;  
 RT "Structures of the human and mouse growth inhibitory factor-encoding  
 RT genes.";  
 RL GENE 144:283-287(1994).  
 CC -1- FUNCTION: INHIBITS SURVIVAL AND NEURITE FORMATION OF CORTICAL  
 CC NEURONS IN VITRO.  
 CC -1- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND  
 CC ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M93310; G199134; -;  
 DR EMBL; S72046; G565192; -;  
 DR PIR; A46034; A46034.  
 DR MGD; MGI:97173; MT3.  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSP; P18055; IMRB.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 30 BETA.  
 FT METAL 31 68 ALPHA.  
 FT METAL 6 6 CLUSTER B (BY SIMILARITY).  
 FT METAL 8 8 CLUSTER B (BY SIMILARITY).  
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).  
 FT METAL 16 16 CLUSTER B (BY SIMILARITY).  
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).  
 FT METAL 22 22 CLUSTER B (BY SIMILARITY).  
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).  
 FT METAL 27 27 CLUSTER B (BY SIMILARITY).  
 FT METAL 30 30 CLUSTER A (BY SIMILARITY).  
 FT METAL 34 34 CLUSTER A (BY SIMILARITY).  
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).  
 FT METAL 37 37 CLUSTER A (BY SIMILARITY).  
 FT METAL 38 38 CLUSTER A (BY SIMILARITY).  
 FT METAL 42 42 CLUSTER A (BY SIMILARITY).  
 FT METAL 45 45 CLUSTER A (BY SIMILARITY).  
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).  
 FT METAL 51 51 CLUSTER A (BY SIMILARITY).  
 FT METAL 64 64 CLUSTER A (BY SIMILARITY).  
 FT METAL 66 66 CLUSTER A (BY SIMILARITY).  
 FT METAL 67 67 CLUSTER A (BY SIMILARITY).  
 SQ SEQUENCE 68 AA; 7009 MW; F15C282F CRC32;

Query Match 65.28; Score 51.5; DB 1; Length 68;  
 Best Local Similarity 75.0%; Pred. No. 0.19;  
 Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 CKCKCKCK-CKC 11  
 | | | | | | | | | |

Db 16 CSDKCKCKGCKC 27  
 RESULT 3  
 MT3\_RAT  
 ID MT3\_RAT STANDARD; PRT; 66 AA.  
 AC P37361;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-III (MT-III) (GROWTH INHIBITORY FACTOR) (GIF).  
 GN MT3.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94018480.  
 RA KOBAYASHI H., UCHIDA Y., IHARA Y., NAKAJIMA K., KOHSAKA S.,  
 RA MIYATAKE T., TSUJI S.;  
 RT "Molecular cloning of rat growth inhibitory factor cDNA and the  
 RT expression in the central nervous system.";  
 RL BRAIN RES. MOL. BRAIN RES. 19:188-194(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=GLIAL TUMOR;  
 RA AMOUREUX M.C., REYSAUS E., WURCH T., COLPAERT F.C., PAUWELS P.J.;  
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND  
 CC ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: BRAIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; S65838; G425382; -;  
 DR EMBL; X89603; G908881; -;  
 DR EMBL; Y08235; E275133; -;  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSP; P18055; IMRB.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 30 BETA.  
 FT METAL 31 66 ALPHA.  
 FT METAL 6 6 CLUSTER B (BY SIMILARITY).  
 FT METAL 8 8 CLUSTER B (BY SIMILARITY).  
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).  
 FT METAL 16 16 CLUSTER B (BY SIMILARITY).  
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).  
 FT METAL 22 22 CLUSTER B (BY SIMILARITY).  
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).  
 FT METAL 27 27 CLUSTER B (BY SIMILARITY).  
 FT METAL 30 30 CLUSTER B (BY SIMILARITY).  
 FT METAL 34 34 CLUSTER A (BY SIMILARITY).  
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).  
 FT METAL 37 37 CLUSTER A (BY SIMILARITY).  
 FT METAL 38 38 CLUSTER A (BY SIMILARITY).  
 FT METAL 42 42 CLUSTER A (BY SIMILARITY).  
 FT METAL 45 45 CLUSTER A (BY SIMILARITY).  
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).



```

FT METAL          51          CLUSTER A (BY SIMILARITY).
FT METAL          62          CLUSTER A (BY SIMILARITY).
FT METAL          64          CLUSTER A (BY SIMILARITY).
FT METAL          65          CLUSTER A (BY SIMILARITY).
SQ SEQUENCE 66 AA; 6809 MW; C03AAC34 CRC32;

Query Match          65.2%; Score 51.5; DB 1; Length 66;
Best Local Similarity 75.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CRCKCKCK-CKC 11
Db 16 CSDCKCKCKCKC 27

RESULT 4
ERG2_NEUCR          STANDARD; PRT; 256 AA.
AC Q92254;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE C-8 STEROL ISOMERASE (DELTA-8--DELTA-7 STEROL ISOMERASE).
GN ERG-1.
OS NEUROSPORA CRASSA.
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;
OC SORDARIALES; SORDARIACEAE; NEUROSPORA.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RA GILBERT J., ORBACH M.J.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CATALYZES THE REACTION WHICH RESULTS IN UNSATURATION AT
CC C-7 IN THE B RING OF STEROLS.
CC -!- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U59671; G1575320;
KW STEROL BIOSYNTHESIS; ISOMERASE; ENDOPLASMIC RETICULUM.
SQ SEQUENCE 256 AA; 28357 MW; 7F7D5EFD CRC32;

Query Match          64.6%; Score 51; DB 1; Length 256;
Best Local Similarity 55.6%; Pred. No. 0.53;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRCKCKCKC 9
Db 30 CRCRCRCRC 38

RESULT 5
MTL_CAEEL          STANDARD; PRT; 74 AA.
ID MTL_CAEEL
AC P17511;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE METALLOTHIONEIN-1 (MT-1).
GN MTL-1 OR MET-1 OR K11G9.6.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTERA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 90094407.
RA SLICE L.W., FREEDMAN J.H., RUBIN C.S.;
RT "Purification, characterization, and cDNA cloning of a novel
RT metallothionein-like, cadmium-binding protein from Caenorhabditis
RT elegans."
RL J. BIOL. CHEM. 265:256-263(1990).
RN [2]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-24.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 90262552.
RA IMAGAWA M., ONOZAWA T., OKUMURA K., OSADA S., NISHIHARA T., KONDO M.;
RT "Characterization of metallothionein cDNAs induced by cadmium in the
RT nematode Caenorhabditis elegans."
RL BIOCHEM. J. 268:237-240(1990).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE; 93155063.
RA FREEDMAN J.H., SLICE L.W., DIXON D., FIRE A., RUBIN C.S.;
RT "The novel metallothionein genes of Caenorhabditis elegans.
RT Structural organization and inducible, cell-specific expression."
RL J. BIOL. CHEM. 268:2554-2564(1993).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 95151184.
RA KUGAWA F., YAMAMOTO H., OSADA S., AOKI M., IMAGAWA M., NISHIHARA T.;
RT "Metallothionein genes in the nematode Caenorhabditis elegans and
RT metal inducibility in mammalian culture cells."
RL BIOMED. ENVIRON. SCI. 7:222-231(1994).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA GRECO T., BRADSHAW H.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: THIS PROTEIN BINDS CATIONS OF SEVERAL TRANSITION
CC ELEMENTS.
CC -!- INDUCTION: BY CADMIUM.
CC -!- DOMAIN: ALL CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS. THESE
CC ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER
CC METALLOTHIONEINS.
CC -!- SIMILARITY: BELONGS TO FAMILY 6 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92909; G156366;
DR EMBL; X53244; G6780;
DR EMBL; D17364; G435621;
DR EMBL; U64853; G1465815;
DR PIR; S09714; S09714.
DR PIR; B45206; B45206.
DR WORMPEP; K11G9.6; CE07379.
DR HSSP; P04355; 4MT2.
KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; CADMIUM.
FT INIT_MET 0
FT CONFLICT 22 22 D -> G (IN AA SEQUENCE).
SQ SEQUENCE 74 AA; 7959 MW; C42B373E CRC32;

Query Match          63.9%; Score 50.5; DB 1; Length 74;
Best Local Similarity 75.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CRCKCKCK-CKC 11
Db 2 CKCKCKCKCKC 13

```

```

RESULT 6
MT21_MYTD
AC P80251; STANDARD; PRT; 71 AA.
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE METALLOTHIONEIN 20-1 ISOFORMS A AND B (MT-20-1A AND MT-20-1B).
OS MYTILUS EDULIS (BLUE MUSSEL).
OC EUKARYOTA; METAZOA; MOLLUSCA; BIVALVIA; PTERIOMORPHIA; MYTILOIDA;
OC MYTILOIDA; MYTILUS.
RN [1]
RP SEQUENCE.
RX MEDLINE; 94062828.
RA MACRAY E.A., OVERNELL J., DUNBAR B., DAVIDSON I., HUNZIKER P.E.,
RA KAEGI J.H.R., FOTHERGILL J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL EUR. J. BIOCHEM. 218:183-194(1993).
CC -!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- INDUCTION: BY CADMIUM.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 IN METALLOTHIONEIN SUPERFAMILY.
DR PIR; S39420; S39420.
DR PFAM; PF00131; metalthio; 1.
DR HSP; P18055; 1MRB.
KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; CADMIUM.
FT VARIANT 66
FT SEQUENCE 71 AA; 6944 MW; 840AA5FA CRC32;
SQ

Query Match 63.3%; Score 50; DB 1; Length 71;
Best Local Similarity 61.5%; Pred. No. 0.31;
Matches 8; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1 CKCK--CKCKKC 11
   ||| ||| |
Db 44 CKCGTCKGCGDC 56

RESULT 7
G3PT_MOUSE
ID G3PT_MOUSE STANDARD; PRT; 440 AA.
AC Q6467; Q60650;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, TESTIS-SPECIFIC
DE (EC 1.2.1.12) (GAPDH).
GN GAPDS OR GAPD-S.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=TESTIS;
RX MEDLINE; 92273722.
RA WELCH J.E., SCHATTE E.C., O'BRIEN D.A., EDDY E.M.;
RT "Expression of a glyceraldehyde 3-phosphate dehydrogenase gene
RT specific to mouse spermatogenic cells.";
RL BIOL. REPROD. 46:869-878(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR SWISS; TISSUE=TESTIS;
RX MEDLINE; 93254745.
RA WELCH J.E., BROWN P.R., O'BRIEN D.A., EDDY E.M.;
RT "Genomic organization of a mouse glyceraldehyde 3-phosphate
RT dehydrogenase gene (gapd-s) expressed in post-meiotic spermatogenic
RT cells.";
RL DEV. GENET. 16:179-189(1995).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING THE SWITCH

```

```

CC BETWEEN DIFFERENT PATHWAYS FOR ENERGY PRODUCTION DURING
CC SPERMIOGENESIS AND IN THE SPERMATOZOON.
CC -!- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + ORTHOPHOSPHATE
CC + NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -!- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: TESTIS-SPECIFIC.
CC -!- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT DAY 20 IN POST-MEIOIC
CC GERM CELLS. LEVELS INCREASE UNTIL DAY 24 AND THEN REMAIN CONSTANT
CC DURING MATURITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M60978; G567204; -.
CC EMBL; U09964; G497413; -.
CC MGD; MGI:95653; GAPDS.
CC PROSITE; PS00071; GAPDH; 1.
CC PFAM; PF00044; gpdh; 1.
CC KW GLYCOLYSIS; OXIDOREDUCTASE; NAD.
CC FT DOMAIN 21 41
CC FT DOMAIN 54 73
CC FT DOMAIN 84 100
CC FT BINDING 256 256
CC FT ACT_SITE 283 283
CC FT CONFLICT 33 34
CC FT CONFLICT 43 43
CC FT SEQUENCE 440 AA; 47657 MW; 40693A26 CRC32;
SQ

Query Match 62.0%; Score 49; DB 1; Length 440;
Best Local Similarity 54.5%; Pred. No. 1.4;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11
   ||| ||| |
Db 21 CPCPCPCPCPC 31

RESULT 8
MT1B_HUMAN
ID MT1B_HUMAN STANDARD; PRT; 61 AA.
AC P07438;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE METALLOTHIONEIN-1B (MT-1B).
GN MT1B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87064506.
RA HEGUY A., WEST A., RICHARDS R.I., KARIN M.;
RT "Structure and tissue-specific expression of the human
RT metallothionein 1B gene.";
RL MOL. CELL. BIOL. 6:2149-2157(1986).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
CC GLUCOCORTICOIDS.
CC -!- SUBUNIT: MONOMER.
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA

```

CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13485; G386962; -  
 DR EMBL; M13484; G386962; JOINED.  
 DR PIR; A25244; A25244.  
 DR MIM; 156349; -  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSSP; P04355; 2MRT.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW CADMIUM; ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 29 BETA.  
 FT METAL 30 61 CLUSTER A.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 7 7 CLUSTER B.  
 FT METAL 13 13 CLUSTER B.  
 FT METAL 15 15 CLUSTER B.  
 FT METAL 19 19 CLUSTER B.  
 FT METAL 21 21 CLUSTER B.  
 FT METAL 24 24 CLUSTER B.  
 FT METAL 26 26 CLUSTER B.  
 FT METAL 29 29 CLUSTER B.  
 FT METAL 33 33 CLUSTER A.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 36 36 CLUSTER A.  
 FT METAL 37 37 CLUSTER A.  
 FT METAL 41 41 CLUSTER A.  
 FT METAL 44 44 CLUSTER A.  
 FT METAL 48 48 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.  
 FT METAL 59 59 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 SQ SEQUENCE 61 AA; 6115 MW; 176D182B CRC32;

Query Match 61.4%; Score 48.5; DB 1; Length 61;  
 Best Local Similarity 62.5%; Pred. No. 0.44;  
 Matches 10; Conservative 0; Mismatches 1; Indels 5; Gaps 2;  
 QY 1 CKCK-CKC-----KCKC 11  
 ||||| |||  
 Db 19 CKCKECKTCKRCKC 34

RESULT 9  
 MT2\_CALSI  
 ID MT2\_CALSI STANDARD; PRT; 58 AA.  
 AC P55950;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-II (MT-II) (MT-IIIB/MT-IIIA).  
 OS CALLINECTES SAPIDUS (BLUE CRAB).  
 OC EUMALACOSTRACA; ARTHROPODA; CRUSTACEA; MALACOSTRACA;  
 OC EUMALACOSTRACA; EUCAIDA; DECAPODA; PLEOCYENATA; BRACHYURA;  
 OC BRACHYRHYNCHA; PORTUNOIDEA; PORTUNIDAE; CALLINECTES.  
 [1]  
 RN SEQUENCE.  
 RX MEDLINE; 96033062.  
 RA BROUWER M., ENGHILD J., HOEXUM-BROUWER T., THOGERSEN I., TRUNCALI A.;

RT "Primary structure and tissue-specific expression of blue crab  
 RT (Callinectes sapidus) metallothionein isoforms.";  
 RL BIOCHEM. J. 311:617-622(1995).  
 CC -!- FUNCTION: BINDS SIX DIVALENT METAL IONS. KNOWN TO BIND COPPER AND  
 CC CADMIUM.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 3 IN METALLOTHIONEIN SUPERFAMILY.  
 CC PFAM; PF00131; metalthio; 1.  
 DR HSSP; P55949; 1DMD.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; COPPER; CADMIUM.  
 FT DOMAIN 1 29 BETA.  
 FT METAL 30 58 ALPHA.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 6 6 CLUSTER B.  
 FT METAL 10 10 CLUSTER B.  
 FT METAL 12 12 CLUSTER B.  
 FT METAL 17 17 CLUSTER B.  
 FT METAL 21 21 CLUSTER B.  
 FT METAL 23 23 CLUSTER B.  
 FT METAL 26 26 CLUSTER B.  
 FT METAL 28 28 CLUSTER B.  
 FT METAL 31 31 CLUSTER A.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 38 38 CLUSTER A.  
 FT METAL 40 40 CLUSTER A.  
 FT METAL 46 46 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 54 54 CLUSTER A.  
 FT METAL 56 56 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.  
 FT VARIANT 1 1 MISSING (IN MT-IIA).  
 SQ SEQUENCE 58 AA; 6287 MW; 097C8AA0 CRC32;

Query Match 61.4%; Score 48.5; DB 1; Length 58;  
 Best Local Similarity 66.7%; Pred. No. 0.42;  
 Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 CKCKCKCK-CKC 11  
 || ||||| |||  
 Db 17 CKTGCKCKSCRC 28

RESULT 10  
 MT1A\_RABIT  
 ID MT1A\_RABIT STANDARD; PRT; 61 AA.  
 AC P11957;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-IA (MT-1A).  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-NEW ZEALAND WHITE;  
 RX MEDLINE; 88240402.  
 RA TAM Y.C., CHOPRA A., HASSAN M., THIRION J.-P.;  
 RT "Cloning, nucleotide sequence and characterization of a New Zealand  
 RT rabbit metallothionein-I gene.";  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 153:209-216(1988).  
 [2]  
 RN SEQUENCE.  
 RP STRAIN-NEW ZEALAND WHITE; TISSUE=LIVER, AND KIDNEY;  
 RC MEDLINE; 95169065.  
 RA HUNZIKER P.E., KAUR P., WAN M., KAENZIG A.;  
 RT "Primary structures of seven metallothioneins from rabbit tissue.";  
 RL BIOCHEM. J. 306:265-270(1995).  
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS.  
 CC -!- SUBUNIT: MONOMER.

CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M20015; G164730; -  
 DR PIR; A27652; A27652  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSSP; P18055; 1MRB.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW CADMIUM; ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT DOMAIN 1 29 BETA.  
 FT DOMAIN 30 61 ALPHA.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 7 7 CLUSTER B.  
 FT METAL 13 13 CLUSTER B.  
 FT METAL 15 15 CLUSTER B.  
 FT METAL 19 19 CLUSTER B.  
 FT METAL 21 21 CLUSTER B.  
 FT METAL 24 24 CLUSTER B.  
 FT METAL 26 26 CLUSTER B.  
 FT METAL 29 29 CLUSTER B.  
 FT METAL 33 33 CLUSTER A.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 36 36 CLUSTER A.  
 FT METAL 37 37 CLUSTER A.  
 FT METAL 41 41 CLUSTER A.  
 FT METAL 44 44 CLUSTER A.  
 FT METAL 48 48 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.  
 FT METAL 59 59 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 SQ SEQUENCE 61 AA; 6103 MW; 9814462B CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
 Best Local Similarity 66.7%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
 Db 15 CASSCKCKCKC 26

RESULT 11  
 MTIG\_HUMAN STANDARD; PRT; 61 AA.  
 AC P13640;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-IG (MT-1G).  
 GN MTIG.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88298811.  
 RA FOSTER R., JAHROUDI N., VARSHNEY U., GEDAMU L.;

RT "Structure and expression of the human metallothionein-IG gene.  
 RT Differential promoter activity of two linked metallothionein-I genes  
 RT in response to heavy metals.";  
 RL J. BIOL. CHEM. 263:11528-11535(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88029905.  
 RA GEDAMU L., VARSHNEY U., JAHROUDI N., FOSTER R., SHWOKAK N.W.;  
 RT "Structure and expression of the human metallothionein genes.";  
 RL EXPERIENTIA SUPPL. 52:361-372(1987).  
 RN [3]  
 RP SEQUENCE OF 1-31 FROM N.A.  
 RX MEDLINE; 94164148.  
 RA PAWELS M., VAN WEYNBERGH J., SOUMILLION A., PROOST P., LEY M.;  
 RT "Induction by zinc of specific metallothionein isoforms in human  
 RT monocytes.";  
 RL EUR. J. BIOCHEM. 220:105-110(1994).  
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03910; G188713; -  
 DR EMBL; S68954; G545719; -  
 DR PIR; A29236; A29236.  
 DR PIR; B61560; B61560.  
 DR MIM; 136353; -  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSSP; P02795; 1MHU.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW CADMIUM; ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 29 BETA.  
 FT DOMAIN 30 61 ALPHA.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 7 7 CLUSTER B.  
 FT METAL 13 13 CLUSTER B.  
 FT METAL 15 15 CLUSTER B.  
 FT METAL 19 19 CLUSTER B.  
 FT METAL 21 21 CLUSTER B.  
 FT METAL 24 24 CLUSTER B.  
 FT METAL 26 26 CLUSTER B.  
 FT METAL 29 29 CLUSTER B.  
 FT METAL 33 33 CLUSTER A.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 36 36 CLUSTER A.  
 FT METAL 37 37 CLUSTER A.  
 FT METAL 41 41 CLUSTER A.  
 FT METAL 44 44 CLUSTER A.  
 FT METAL 48 48 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.  
 FT METAL 59 59 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 SQ SEQUENCE 61 AA; 6070 MW; AED6C071 CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
 Best Local Similarity 66.7%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
 | | | | | | |  
 DB 15 CASSCKCKCKC 26

RESULT 12  
 ID MTIR\_HUMAN STANDARD; PRT; 62 AA.  
 AC P80236;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-1K (MT-1K).  
 GN MTIK  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=LIVER;  
 RA HONZIKER P.E.;  
 RL SUBMITTED (NOV-1993) TO THE SWISS-PROT DATA BANK.

CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC MIM: 156357; -

DR PROSITE; P500203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.  
 DR HSP; P02795; 1MHU  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW CADMIUM; ACETYLATION.

FT MOD\_RES 1 1 ACETYLATION.  
 FT DOMAIN 1 30 BETA.  
 FT DOMAIN 31 62 ALPHA.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 7 7 CLUSTER B.  
 FT METAL 14 14 CLUSTER B.  
 FT METAL 16 16 CLUSTER B.  
 FT METAL 20 20 CLUSTER B.  
 FT METAL 22 22 CLUSTER B.  
 FT METAL 25 25 CLUSTER B.  
 FT METAL 27 27 CLUSTER B.  
 FT METAL 30 30 CLUSTER B.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 35 35 CLUSTER A.  
 FT METAL 37 37 CLUSTER A.  
 FT METAL 38 38 CLUSTER A.  
 FT METAL 42 42 CLUSTER A.  
 FT METAL 45 45 CLUSTER A.  
 FT METAL 49 49 CLUSTER A.  
 FT METAL 51 51 CLUSTER A.  
 FT METAL 58 58 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 FT METAL 61 61 CLUSTER A.  
 SQ SEQUENCE 62 AA; 6141 MW; 1F0973C0 CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 62;  
 Best Local Similarity 66.7%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
 | | | | | | |  
 DB 16 CASSCKCKCKC 27

RESULT 13  
 ID MTIR\_HUMAN STANDARD; PRT; 61 AA.  
 AC Q93083;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-IR (MT-IR).  
 GN MTIR.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=RETICULOCYTES;  
 RX MEDLINE; 96338109.

RA LAMBERT E., KILLE P., SWAMINATHAN R.;  
 RT "Cloning and sequencing a novel metallothionein I isoform expressed  
 in human reticulocytes";  
 RL FEBS LETT. 389:210-212(1996).  
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS (BY SIMILARITY).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN RETICULOCYTES.  
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X97261; E239986; -  
 CC PROSITE; P500203; METALLOTHIONEIN\_VRT; 1.  
 CC PFAM; PF00131; metalthio; 1.  
 CC HSP; P18055; 1MRB.  
 KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;  
 KW CADMIUM; ACETYLATION.  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 1 29 BETA.  
 FT DOMAIN 30 61 ALPHA.  
 FT METAL 5 5 CLUSTER B.  
 FT METAL 7 7 CLUSTER B.  
 FT METAL 13 13 CLUSTER B.  
 FT METAL 15 15 CLUSTER B.  
 FT METAL 19 19 CLUSTER B.  
 FT METAL 21 21 CLUSTER B.  
 FT METAL 24 24 CLUSTER B.  
 FT METAL 26 26 CLUSTER B.  
 FT METAL 29 29 CLUSTER B.  
 FT METAL 33 33 CLUSTER A.  
 FT METAL 34 34 CLUSTER A.  
 FT METAL 36 36 CLUSTER A.  
 FT METAL 37 37 CLUSTER A.  
 FT METAL 41 41 CLUSTER A.  
 FT METAL 44 44 CLUSTER A.  
 FT METAL 48 48 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.

FT METAL 59 59 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 SQ SEQUENCE 61 AA; 6062 MW; F6D451A5 CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
 Best Local Similarity 66.7%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
 | | | | | | | |  
 DB 15 CASSCKCKCKC 26

## RESULT 14

ID MT2B\_RABIT STANDARD; PRT; 61 AA.  
 AC P80289;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-IIB (MT-2B).  
 OS ERYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=LIVER, AND KIDNEY;  
 RX MEDLINE: 92140112.  
 RA HUNZIKER P.E.;  
 RT "Amino acid sequence determination."  
 RL METH. ENZYMOLOGY. 205:421-426(1991).  
 RN [2]

RP SEQUENCE.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=LIVER, AND KIDNEY;  
 RX MEDLINE: 95169065.

RA HUNZIKER P.E.; KAUR P., WAN M., KAENZIG A.;  
 RT "Primary structures of seven metallothioneins from rabbit tissue."  
 RL BIOCHEM. J. 306:265-270(1995).  
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS.

CC -!- SUBUNIT: MONOMER.

CC -!- DOMAIN: CLASS 1 METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC PROSITE; P500203; METALLOTHIONEIN\_VRT; 1.  
 DR PFAM; PF00131; metalthio; 1.

DR HSSP; P18055; IMRB.

KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;

KW CADMIUM; ACETYLATION.

FT MOD\_RES 1 1 ACETYLATION.

FT DOMAIN 1 29 BETA.

FT DOMAIN 30 61 ALPHA.

FT METAL 5 5 CLUSTER B.

FT METAL 7 7 CLUSTER B.

FT METAL 13 13 CLUSTER B.

FT METAL 15 15 CLUSTER B.

FT METAL 19 19 CLUSTER B.

FT METAL 21 21 CLUSTER B.

FT METAL 24 24 CLUSTER B.

FT METAL 26 26 CLUSTER B.

FT METAL 29 29 CLUSTER B.

FT METAL 33 33 CLUSTER A.

FT METAL 34 34 CLUSTER A.

FT METAL 36 36 CLUSTER A.

FT METAL 37 37 CLUSTER A.

FT METAL 41 41 CLUSTER A.

FT METAL 44 44 CLUSTER A.

FT METAL 48 48 CLUSTER A.  
 FT METAL 50 50 CLUSTER A.  
 FT METAL 57 57 CLUSTER A.  
 FT METAL 59 59 CLUSTER A.  
 FT METAL 60 60 CLUSTER A.  
 SQ SEQUENCE 61 AA; 6104 MW; 305A2491 CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
 Best Local Similarity 66.7%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
 | | | | | | | |  
 DB 15 CASSCKCKCKC 26

## RESULT 15

ID MT2D\_RABIT STANDARD; PRT; 61 AA.  
 AC P80291;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE METALLOTHIONEIN-IID (MT-2D).  
 OS ERYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=LIVER, AND KIDNEY;  
 RX MEDLINE: 95169065.  
 RA HUNZIKER P.E.; KAUR P., WAN M., KAENZIG A.;  
 RT "Primary structures of seven metallothioneins from rabbit tissue."  
 RL BIOCHEM. J. 306:265-270(1995).  
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE  
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND  
 CC GLUCOCORTICOIDS.

CC -!- SUBUNIT: MONOMER.

CC -!- DOMAIN: CLASS 1 METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 CC PROSITE; P500203; METALLOTHIONEIN\_VRT; 1.

DR PFAM; PF00131; metalthio; 1.

DR HSSP; P18055; IMRB.

KW METAL-BINDING; METAL-THIOLATE CLUSTER; CHELATION; ZINC; COPPER;

KW CADMIUM; ACETYLATION.

FT MOD\_RES 1 1 ACETYLATION.

FT DOMAIN 1 29 BETA.

FT DOMAIN 30 61 ALPHA.

FT METAL 5 5 CLUSTER B.

FT METAL 7 7 CLUSTER B.

FT METAL 13 13 CLUSTER B.

FT METAL 15 15 CLUSTER B.

FT METAL 19 19 CLUSTER B.

FT METAL 21 21 CLUSTER B.

FT METAL 24 24 CLUSTER B.

FT METAL 26 26 CLUSTER B.

FT METAL 29 29 CLUSTER B.

FT METAL 33 33 CLUSTER A.

FT METAL 34 34 CLUSTER A.

FT METAL 36 36 CLUSTER A.

FT METAL 37 37 CLUSTER A.

FT METAL 41 41 CLUSTER A.

FT METAL 44 44 CLUSTER A.

FT METAL 50 50 CLUSTER A.

FT METAL 57 57 CLUSTER A.

FT METAL 59 59 CLUSTER A.

FT METAL 60 60 CLUSTER A.  
SQ SEQUENCE 61 AA: 6173 MW: A9246F92 CRC32;

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
Best Local Similarity 66.7%; Pred. No. 0.59;  
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
| | | | |  
Db 15 CASSCKCKCKC 26

Search completed: September 7, 1999, 23:50:16  
Job time: 1955 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:28 ; Search time 116.8 Seconds  
(without alignments)  
3.773 Million cell updates/sec

Title: US-09-124-280A-13  
Perfect score: 79  
Sequence: 1 CKCKCKCKCK 11

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	68.4	160	JQ0542	185K secretory pro
2	53.5	67.7	1700	S08167	Balbani ring 3 pr
3	51.5	65.2	68	A46034	metallothionein 3,
4	51.5	65.2	66	S58086	metallothionein 3
5	51.5	65.2	68	I67866	growth inhibitory
6	50.5	63.9	75	B45206	metallothionein 2
7	50	63.3	71	S39420	metallothionein 20
8	50	63.3	71	S47576	metallothionein 20
9	49	62.0	440	I49681	glyceroldehyde-3-p
10	49	62.0	618	T00476	probable vacuolar
11	49	62.0	248	E71602	probable integral
12	48.5	61.4	61	SMHU1B	metallothionein 1B
13	48.5	61.4	57	S59073	metallothionein 1B
14	47.5	60.1	61	SMRT2	metallothionein II
15	47.5	60.1	61	SMHY2C	metallothionein II
16	47.5	60.1	61	SMHU1G	metallothionein IG
17	47.5	60.1	61	S69277	metallothionein 1R
18	47.5	60.1	61	A27652	metallothionein 1A
19	47.5	60.1	61	S54334	metallothionein-2B
20	47.5	60.1	61	S54332	metallothionein-2B
21	47.5	60.1	61	S54333	metallothionein-2E
22	47.5	60.1	61	I48173	metallothionein II
23	47.5	60.1	61	I48116	metallothionein II
24	47.5	60.1	61	S18403	metallothionein II
25	47.5	60.1	58	S43367	metallothionein -
26	47.5	60.1	75	S17156	metallothionein -
27	47	59.5	407	EDBEQ3	Immediate-early pr
28	47	59.5	303	S40973	hypothetical prote
29	46.5	58.9	61	SMHU2	metallothionein 2
30	46.5	58.9	61	SMK2	metallothionein 2
31	46.5	58.9	61	SMHU1E	metallothionein 1E
32	46.5	58.9	61	SMHU1A	metallothionein 1A
33	46.5	58.9	61	SMK1	metallothionein 1
34	46.5	58.9	61	SMHU1F	metallothionein 1F
35	46.5	58.9	61	SMMS2	metallothionein II
36	46.5	58.9	61	S00811	metallothionein II
37	46.5	58.9	63	S08190	metallothionein 1
38	46.5	58.9	61	B23889	metallothionein 2
39	46.5	58.9	61	S47652	metallothionein 1X

40 46.5 58.9 61 2 S47651 metallothionein 1H  
41 46.5 58.9 61 2 I57572 metallothionein II  
42 46.5 58.9 72 2 S39419 metallothionein 10  
43 46.5 58.9 72 2 S39416 metallothionein 10  
44 46.5 58.9 72 2 S39417 metallothionein 10  
45 46.5 58.9 72 2 S39418 metallothionein 10

ALIGNMENTS

RESULT 1

JQ0542  
185K secretory protein - midge (Chironomus tentans) (fragment)  
N:Alternate names: balbiani ring 3 protein  
C:Species: Chironomus tentans  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 24-Feb-1995  
C:Accession: JQ0542  
R:Dignam, S.S.; Case, S.T.  
Gene 88, 133-140, 1990  
A:Title: Balbiani ring 3 in Chironomus tentans encodes a 185-kDa secretory protein wh  
A:Reference number: JQ0542; MUID:90269600  
A:Accession: JQ0542  
A:Molecule type: mRNA  
A:Residues: 1-160 <DIG>  
A:Cross-references: GB:M24160  
A:Experimental source: salivary gland

Query Match 68.4%; Score 54; DB 2; Length 160;  
Best Local Similarity 36.0%; Pred. No. 0.51;  
Matches 9; Conservative 2; Mismatches 0; Indels 14; Gaps 1;

Qy 1 CKCKCK-----CKCKC 11  
Db 58 CECKCKNNPSTSPQVWDADCECKC 82

RESULT 2

S08167  
Balbiani ring 3 protein - midge (Chironomus tentans)  
C:Species: Chironomus tentans  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Mar-1998  
C:Accession: S08167  
R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.  
J. Mol. Biol. 211, 331-349, 1990  
A:Title: The balbiani ring 3 gene in Chironomus tentans has a diverged repetitive str  
A:Reference number: S08167; MUID:90172404  
A:Accession: S08167  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-1700 <PAU>  
A:Cross-references: GB:X52263; NID:g7057; PID:g7058  
C:Genetics:  
A:Gene: BR3  
A:Map position: 4

Query Match 67.7%; Score 53.5; DB 2; Length 1700;  
Best Local Similarity 34.6%; Pred. No. 2.5;  
Matches 9; Conservative 2; Mismatches 0; Indels 15; Gaps 1;

Qy 1 CKCKCK-----CKCKC 11  
Db 589 CECKCKNNPCTSPQVWDADCECKC 614

RESULT 3

A46034  
metallothionein 3, brain-specific - mouse  
N:Alternate names: neurotrophic growth inhibitory factor  
C:Species: Mus musculus (house mouse)  
C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 24-Apr-1998

C:Accession: A46034  
R:Palmiter, R.D.; Findley, S.D.; Whitmore, T.E.; Durnam, D.M.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6333-6337, 1992  
A:Title: MT-III, a brain-specific member of the metallothionein gene family.  
A:Reference number: A46034; MUID:92335292  
A:Accession: A46034  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-68 <PAL>  
A:CROSS-references: GB:M93310; NID:g199133; PID:g199134  
A:Note: sequence extracted from NCBI backbone (NCBIN:108715, NCBIN:111115, NCBIP:108716)  
C:Superfamily: metallothionein

Query Match 65.2%; Score 51.5; DB 2; Length 68;  
Best Local Similarity 75.0%; Pred. No. 0.62;  
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
| | | | | | | |  
Db 16 CSDKCKCKGCKC 27

RESULT 4  
S58086  
metallothionein 3 - rat  
N:Alternate names: neurotrophic growth inhibitory factor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 24-Apr-1998  
C:Accession: S58086; I52636  
R:Amoureux, M.C.; Rethaus, E.; Wurch, T.; Colpaert, F.C.; Pauwels, P.J.  
submitted to the EMBL Data Library, July 1995  
A:Reference number: S58084  
A:Accession: S58086  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-66 <AMO>  
A:CROSS-references: EMBL:X89603; NID:g908880; PID:g908881  
R:Kobayashi, H.; Uchida, Y.; Ihara, Y.; Nakajima, K.; Kohsaka, S.; Miyatake, T.; Tsuji,  
Brain Res. Mol. Brain Res. 19, 188-194, 1993  
A:Title: Molecular cloning of rat growth inhibitory factor cDNA and the expression in th  
A:Reference number: I52636; MUID:94018480  
A:Accession: I52636  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-66 <KOB>  
A:CROSS-references: GB:S65838; NID:g425381; PID:g425382  
C:Superfamily: metallothionein  
C:Keywords: acetylated amino end; chelation; metal binding; metal-thiolate cluster  
F:1/Modified site: acetylated amino end (Met) #status predicted  
F:6,14,16,20,22,25,27,30/Binding site: transition metal ions (Cys) #status predicted  
F:34,35,37,38,42,45,49,51,62,64,65/Binding site: transition metal ions (Cys) #status pre

Query Match 65.2%; Score 51.5; DB 2; Length 66;  
Best Local Similarity 75.0%; Pred. No. 0.61;  
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
| | | | | | | |  
Db 16 CSDKCKCKGCKC 27

RESULT 5  
I67866  
growth inhibitory factor - mouse  
C:Species: Mus sp. (mouse)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 02-Jul-1998  
C:Accession: I67866  
R:Naruse, S.; Igarashi, S.; Furuya, T.; Kobayashi, H.; Miyatake, T.; Tsuji, S.  
Gene 144, 283-287, 1994  
A:Title: Structures of the human and mouse growth inhibitory factor-encoding genes.  
A:Reference number: I53803; MUID:94314230

A:Accession: I67866  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-68 <RES>  
A:CROSS-references: GB:S72046; NID:g565191; PID:g565192  
C:Genetics:  
A:Gene: GIF  
A:Introns: 11/1; 33/1  
C:Superfamily: metallothionein

Query Match 65.2%; Score 51.5; DB 2; Length 68;  
Best Local Similarity 75.0%; Pred. No. 0.62;  
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
| | | | | | | |  
Db 16 CSDKCKCKGCKC 27

RESULT 6  
B45206  
metallothionein 2 - Caenorhabditis elegans  
N:Alternate names: metallothionein I  
C:Species: Caenorhabditis elegans  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Sep-1997  
C:Accession: B45206; S09714  
R:Freedman, J.H.; Slice, L.W.; Dixon, D.; Fire, A.; Rubin, C.S.  
J. Biol. Chem. 268, 2554-2564, 1993  
A:Title: The novel metallothionein genes of Caenorhabditis elegans. Structural organi  
A:Reference number: A45206; MUID:93155063  
A:Accession: B45206  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-75 <FRE>  
A:Note: sequence extracted from NCBI backbone (NCBIP:124147)  
R:Imagawa, M.; Onozawa, T.; Okumura, K.; Osada, S.; Nishihara, T.; Kondo, M.  
Biochem. J. 268, 237-240, 1990  
A:Title: Characterization of metallothionein cDNAs induced by cadmium in the nematode  
A:Reference number: S09714; MUID:90262552  
A:Accession: S09714  
A:Molecule type: mRNA  
A:Residues: 1-75 <IMA>  
A:CROSS-references: EMBL:X53244; NID:g6779; PID:g6780  
C:Superfamily: metallothionein

Query Match 63.9%; Score 50.5; DB 2; Length 75;  
Best Local Similarity 75.0%; Pred. No. 0.87;  
Matches 9; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKC 11  
| | | | | | | |  
Db 3 CKCDCKCKGCKC 14

RESULT 7  
S39420  
metallothionein 20-Ia - blue mussel  
C:Species: Mytilus edulis (blue mussel)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Mar-1999  
C:Accession: S39420  
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.  
Eur. J. Biochem. 218, 183-194, 1993  
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of me  
A:Reference number: S39416; MUID:94062828  
A:Accession: S39420  
A:Molecule type: protein  
A:Residues: 1-71 <MAC>  
C:Superfamily: metallothionein  
C:Keywords: chelation; metal binding

Query Match 63.3%; Score 50; DB 2; Length 71;  
 Best Local Similarity 61.5%; Pred. No. 0.97;  
 Matches 8; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 CKCK--CKCKCKC 11  
 ||| ||| |  
 Db 44 CKCSGTCGCGDC 56

## RESULT 8

metallothionein 20-Ib - blue mussel  
 C:Species: Mytilus edulis (blue mussel)  
 C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 17-Mar-1999  
 C:Accession: S47576  
 R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.; F.  
 Eur. J. Biochem. 218, 183-194, 1993  
 A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of metal  
 A:Reference number: S39416; MUID:94062828  
 A:Accession: S47576  
 A:Molecule type: protein  
 A:Residues: 1-71 <MAC>  
 C:Superfamily: metallothionein  
 C:Keywords: Chelation; metal binding

Query Match 63.3%; Score 50; DB 2; Length 71;  
 Best Local Similarity 61.5%; Pred. No. 0.97;  
 Matches 8; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 CKCK--CKCKCKC 11  
 ||| ||| |  
 Db 44 CKCSGTCGCGDC 56

## RESULT 9

glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 13-Mar-1997  
 C:Accession: I49681; I49046  
 R:Welch, J.E.; Schatte, E.C.; O'Brien, D.A.; Eddy, E.M.  
 Biol. Reprod. 46, 869-878, 1992  
 A:Title: Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mouse  
 A:Reference number: I49681; MUID:92273722  
 A:Accession: I49681  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-440 <RES>  
 A:Cross-references: GB:M60978; NID:gl93425; PID:g567204  
 R:Welch, J.E.; Brown, P.R.; O'Brien, D.A.; Eddy, E.M.  
 Dev. Genet. 16, 179-189, 1995  
 A:Title: Genomic organization of a mouse glyceraldehyde 3-phosphate dehydrogenase gene  
 A:Reference number: I49046; MUID:95254745  
 A:Accession: I49046  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-32, 35-42, 'V', 44-440 <RE2>  
 A:Cross-references: EMBL:U09964; NID:g497412; PID:g497413  
 C:Genetics:  
 A:Gene: Gapd-S  
 A:Introns: 35/1; 114/2; 146/3; 182/2; 212/3; 252/2; 279/3; 330/2; 384/3; 417/2  
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase  
 C:Keywords: oxidoreductase

Query Match 62.0%; Score 49; DB 2; Length 440;  
 Best Local Similarity 54.5%; Pred. No. 3.9;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
 ||| ||| |  
 Db 21 CPCPCPCPCPC 31

## RESULT 10

T00476  
 probable vacuolar sorting receptor protein F1913.17 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
 C:Accession: T00476  
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K.  
 submitted to the EMBL Data Library, April 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC F1913 genomic sequence.  
 A:Reference number: Z14153  
 A:Accession: T00476  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-618 <ROU>  
 A:Cross-references: EMBL:AC004238; NID:g3033373; PID:g3033390  
 C:Genetics:  
 A:Map position: II  
 A:Introns: 100/1; 329/1; 358/3; 382/1; 410/1; 435/3; 463/1; 490/3; 512/1; 555/1; 590/  
 A:Note: F1913.17

Query Match 62.0%; Score 49; DB 2; Length 618;  
 Best Local Similarity 52.6%; Pred. No. 4.7;  
 Matches 10; Conservative 0; Mismatches 1; Indels 8; Gaps 2;

QY 1 CKCK-CKCK-----CKC 11  
 ||| ||| ||| |||  
 Db 522 CRDGGCKRNNMGGECKC 540

## RESULT 11

E71602  
 Probable integral membrane protein PFB0950w - malaria parasite Plasmodium falciparum  
 C:Species: Plasmodium falciparum  
 C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
 C:Accession: E71602  
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
 ; Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
 Science 282, 1126-1132, 1998  
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
 A:Reference number: A71600  
 A:Accession: E71602  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-248 <GAR>  
 A:Cross-references: GB:AE001428; GB:AE001362; NID:g3845316; PID:g3845320; TIGR:PFB095  
 A:Experimental source: clone 3D7  
 C:Genetics:  
 A:Gene: PFB0950w

Query Match 62.0%; Score 49; DB 2; Length 248;  
 Best Local Similarity 54.5%; Pred. No. 2.7;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
 ||| ||| |||  
 Db 173 CSCICSCICSC 183

## RESULT 12

SMH1B  
 metallothionein 1B - human  
 C:Species: Homo sapiens (man)  
 C:Date: 25-Oct-1987 #sequence\_revision 09-Sep-1994 #text\_change 18-Sep-1998  
 C:Accession: A25244  
 R:Heguy, A.; West, A.; Richards, R.I.; Karin, M.  
 Mol. Cell. Biol. 6, 2149-2157, 1986  
 A:Title: Structure and tissue-specific expression of the human metallothionein IB gen  
 A:Reference number: A25244; MUID:87064506  
 A:Accession: A25244

A:Molecule type: DNA  
A:Residues: 1-61 <HEG>  
A:Cross-references: GB:M13484; GB:M13485; NID:g188709; PID:g386962  
C:Genetics:  
A:Gene: GDB:MT1B; MT1  
A:Cross-references: GDB:125564; OMIM:156349  
A:Map position: 16q13-16q13  
A:Introns: 10/1; 32/1  
C:Superfamily: metallothionein  
C:Keywords: acetylated amino end; chelation; metal binding; metal-thiolate cluster  
F:1-29/Domain: beta <NH2>  
F:30-61/Domain: alpha <ALP>  
F:1/Modified site: acetylated amino end (Met) #status predicted  
F:5,7,13,15,19,21,24,26,29/Binding site: transition metal ions (Cys) #status predicted  
F:33,34,36,37,41,44,48,50,57,59,60/Binding site: transition metal ions (Cys) #status predicted

Query Match 61.4%; Score 48.5; DB 1; Length 61;  
Best Local Similarity 62.3%; Pred. No. 1.3;  
Matches 10; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

QY 1 CKCK-CRCK-CKCK 11  
| | | | | | | | | |  
Db 19 CKCKCKCKTSCKKCC 34

RESULT 13  
SMRT2  
metallothionein isoform IIA - blue crab  
C:Species: Callinectes sapidus (blue crab)  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 12-Feb-1999  
C:Accession: S59073  
R:Brouwer, M.; Englund, J.; Hoexum-Brouwer, T.; Thogersen, I.; Truncali, A.  
Biochem. J. 311, 617-622, 1995  
A:Title: Primary structure and tissue-specific expression of blue crab (Callinectes sapidus) metallothionein  
A:Reference number: S59072  
A:Accession: S59073  
A:Molecule type: protein  
A:Residues: 1-57 <BRO>  
C:Superfamily: metallothionein  
C:Keywords: metal binding

Query Match 61.4%; Score 48.5; DB 2; Length 57;  
Best Local Similarity 66.7%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 CKCKCKCK-CKCK 11  
| | | | | | | | | |  
Db 16 CKCKCKCKCKCK 27

RESULT 14  
SMRT2  
metallothionein II - rat  
N:Alternate names: MT-II  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 18-Mar-1997  
C:Accession: B61561; S02630; A03274; S65713  
R:Andersen, R.D.; Taplitz, S.J.; Birren, B.W.; Bristol, G.; Herschman, H.R.  
Experientia Suppl. 52, 373-384, 1987  
A:Title: Rat metallothionein multigene family.  
A:Reference number: A61561; MUID:88029906  
A:Accession: B61561  
A:Molecule type: DNA  
A:Residues: 1-61 <AND>  
R:Woergetter, E.; Wagner, G.; Vasak, M.; Kaegi, J.H.R.; Wuethrich, K.  
Eur. J. Biochem. 167, 457-466, 1987  
A:Title: Sequence-specific (1H-NMR assignments in rat-liver metallothionein-2.  
A:Reference number: S02630; MUID:88004467  
A:Accession: S02630  
A:Molecule type: protein  
A:Residues: 1-61 <WOE>

R:Winge, D.R.; Nielson, K.B.; Zeikus, R.D.; Gray, W.R.  
J. Biol. Chem. 259, 11419-11425, 1984  
A:Title: Structural characterization of the isoforms of neonatal and adult rat liver metallothionein  
A:Reference number: A92445; MUID:84289593  
A:Accession: A03274  
A:Molecule type: protein  
A:Residues: 1-61 <WIN>  
A:Experimental source: neonatal and adult liver  
R:Saito, S.; Hunziker, P.E.  
Biochim. Biophys. Acta 1289, 65-70, 1996  
A:Title: Differential sensitivity of metallothionein-1 and -2 in liver of zinc-injected rats  
A:Reference number: S65712; MUID:96195842  
A:Accession: S65713  
A:Molecule type: protein  
A:Residues: 3-61 <SAI>  
C:Comment: The vertebrate metallothioneins contain two metal-binding domains. Cluster C:Superfamily: metallothionein  
C:Keywords: blocked amino end; chelation; metal binding; metal-thiolate cluster  
F:1-29/Domain: beta #status predicted <NH2>  
F:30-61/Domain: alpha #status predicted <ALP>  
F:1/Modified site: blocked amino end (Met) (probably acetylated) #status experimental  
F:5,7,13,15,19,21,24,26,29/Binding site: transition metal ions (Cys) #status predicted  
F:33,34,36,37,41,44,48,50,57,59,60/Binding site: transition metal ions (Cys) #status predicted

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
Best Local Similarity 64.3%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
| | | | | | | | | |  
Db 13 CSCAGSCKCKCKC 26

RESULT 15  
SMRT2  
metallothionein II - Chinese hamster  
C:Species: Cricetus griseus (Chinese hamster)  
C>Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 22-Jan-1999  
C:Accession: A03276; S50199  
R:Griffith, B.B.; Walters, R.A.; Enger, M.D.; Hildebrand, C.E.; Griffith, J.K.  
Nucleic Acids Res. 11, 901-910, 1983  
A:Title: cDNA cloning and nucleotide sequence comparison of Chinese hamster metallothionein  
A:Reference number: A93501; MUID:83168910  
A:Accession: A03276  
A:Molecule type: mRNA  
A:Residues: 1-61 <GRI>  
A:Cross-references: GB:J00062; NID:g191139; PID:g304523  
R:Yamada, K.; Kato, H.; Kanda, N.; Fujii-Kuriyama, Y.; Utakoji, T.; Itoh, R.  
Biochim. Biophys. Acta 1219, 581-591, 1994  
A:Title: Sequence homology of Chinese hamster metallothionein genes I and II to those of other mammals  
A:Reference number: S50198  
A:Accession: S50199  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-61 <YAM>  
C:Superfamily: metallothionein  
C:Keywords: metal binding  
F:5,7,13,15,19,21,24,26,29/Binding site: transition metal ions (Cys) #status predicted  
F:33,34,36,37,41,44,48,50,57,59,60/Binding site: transition metal ions (Cys) #status predicted

Query Match 60.1%; Score 47.5; DB 1; Length 61;  
Best Local Similarity 64.3%; Pred. No. 1.8;  
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
| | | | | | | | | |  
Db 13 CSCAGSCKCKCKC 26

Search completed: September 7, 1999, 23:06:28

Job time: 2476 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:35 ; Search time 80.79 Seconds  
(without alignments)  
1.344 Million cell updates/sec

Title: US-09-124-280A-13  
Perfect score: 79  
Sequence: 1 CKCKCKCKCK 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79	100.0	11	2	US-08-456-112B-13
2	64	81.0	257	1	US-08-414-926A-7
3	48.5	61.4	61	2	US-08-785-530-3
4	47.5	60.1	61	2	US-08-785-530-6
5	47	59.5	801	1	US-07-906-349A-6
6	46.5	58.9	61	2	US-08-785-530-1
7	46.5	58.9	61	2	US-08-785-530-4
8	46.5	58.9	61	2	US-08-785-530-5
9	46.5	58.9	254	2	US-08-787-026-7
10	46	58.2	40	4	5240848-11
11	44	55.7	21	2	US-08-448-418-106
12	43.5	55.1	14	1	US-08-322-962-12
13	43.5	55.1	189	1	US-08-469-427A-15
14	43.5	55.1	231	3	PCT-US96-09001-10
15	43.5	55.1	40	4	5219739-22
16	43.5	55.1	40	4	5240848-7
17	42.5	53.8	68	1	US-07-696-051B-1
18	42.5	53.8	68	1	US-07-924-063A-1
19	42.5	53.8	68	1	US-08-138-340B-2
20	42	53.2	40	4	532871-3
21	41.5	52.5	419	3	PCT-US96-09001-2
22	41	51.9	40	4	532671-4
23	40.5	51.3	39	1	US-08-036-555B-40
24	40.5	51.3	39	1	US-08-469-569-40
25	40.5	51.3	39	1	US-08-249-322A-40
26	40.5	51.3	39	2	US-08-469-526A-40
27	40.5	51.3	39	2	US-08-734-591A-40
28	40.5	51.3	39	2	US-08-469-660-40
29	40.5	51.3	39	3	PCT-US94-05083C-40
30	40.5	51.3	39	3	PCT-US95-06846A-40
31	40	50.6	10	1	US-08-097-830E-16
32	40	50.6	10	2	US-08-456-112B-16
33	39	49.4	362	1	US-08-415-751-6
34	39	49.4	362	1	US-08-415-751-35
35	39	49.4	41	1	US-08-451-947-97
36	39	49.4	11	2	US-08-218-026-50
37	39	49.4	45	2	US-08-424-826A-97
38	39	49.4	11	2	US-08-653-632-50
39	39	49.4	2050	2	US-08-347-594A-2

Sequence 97, Appl  
Sequence 28, Appl  
Sequence 93, Appl  
Sequence 94, Appl  
Sequence 95, Appl  
Sequence 96, Appl

ALIGNMENTS

RESULT 1  
US-08-456-112B-13  
; Sequence 13, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-13  
  
Query Match 100.0%; Score 79; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00017;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 CKCKCKCKCK 11  
Db 1 CKCKCKCKCK 11  
  
RESULT 2  
US-08-414-926A-7  
; Sequence 7, Application US/08414926A  
; Patent No. 5721354  
; GENERAL INFORMATION:  
; APPLICANT: Spaete, Richard  
; APPLICANT: Cha, Tai-An  
; TITLE OF INVENTION: NOVEL HUMAN CYTOMEGALOVIRUS  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum

STREET: 5 Palo Alto Square  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306-2155  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,926A  
FILING DATE: March 31, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cseri, Luann  
REGISTRATION NUMBER: 31,822  
REFERENCE/DOCKET NUMBER: AVIR-011/OOUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-494-7622  
TELEFAX: 415-857-0663  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 257 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IMMEDIATE SOURCE:  
CLONE: tol.01  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..257  
OTHER INFORMATION: /label= UL133  
US-08-414-926A-7

Query Match 81.0%; Score 64; DB 1; Length 257;  
Best Local Similarity 72.7%; Pred. No. 0.078;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CKCKCKCKCK 11  
| | | | | | | | | |  
Db 87 CCGCKCKCKCK 97

RESULT 3  
US-08-785-530-3  
Sequence 3, Application US/08785530  
Patent No. 5814480  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: A NOVEL HUMAN METALLOTHIONINE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/785,530  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0194 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 386962  
US-08-785-530-3

Query Match 61.4%; Score 48.5; DB 2; Length 61;  
Best Local Similarity 62.5%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

Qy 1 CKCK-CKC----KCKC 11  
| | | | | | | | | |  
Db 19 CKCKCKCTCKKCKC 34

RESULT 4  
US-08-785-530-6  
Sequence 6, Application US/08785530  
Patent No. 5814480  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: A NOVEL HUMAN METALLOTHIONINE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/785,530  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0194 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank



```
; CLONE: 305363
US-08-785-530-6

Query Match 60.1%; Score 47.5; DB 2; Length 61;
Best Local Similarity 64.3%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CRC 11
   | | | | | | |
Db 13 CSCSGCKCKCKC 26

RESULT 5
US-07-906-349A-6
; Sequence 6, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; APPLICANT: Skolnik, Edward F.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-349A-6

Query Match 59.5%; Score 47; DB 1; Length 801;
Best Local Similarity 54.5%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11
   | | | | |
Db 318 CACTCCCTCAC 328

RESULT 6
US-08-785-530-1
; Sequence 1, Application US/08785530
; Patent No. 5814480
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.

Query Match 58.9%; Score 46.5; DB 2; Length 61;
Best Local Similarity 64.3%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CRC 11
   | | | | |
Db 13 CACTGCKCKCKC 26

RESULT 7
US-08-785-530-4
; Sequence 4, Application US/08785530
; Patent No. 5814480
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN METALLOTHIONINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,530
```

;; FILING DATE: Herewith  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Billings, Lucy J.  
;; REGISTRATION NUMBER: 36,749  
;; REFERENCE/DOCKET NUMBER: PF-0194 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-855-0555  
;; TELEFAX: 415-845-4166  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 61 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: GenBank  
;; CLONE: 386865  
US-08-785-530-4

Query Match 58.9%; Score 46.5; DB 2; Length 61;  
Best Local Similarity 64.3%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 KCCK--KCCK--CKC 11  
| | | | | | | | | |  
DB 13 CTCAGSCKCKCKC 26

RESULT 8  
US-08-785-530-5  
; Sequence 5, Application US/08785530  
; Patent No. 5814480  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Goli, Surya K.  
; TITLE OF INVENTION: A NOVEL HUMAN METALLOTHIONINE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/785,530  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0194 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 61 amino acids

;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: GenBank  
;; CLONE: 342385  
US-08-785-530-5

Query Match 58.9%; Score 46.5; DB 2; Length 61;  
Best Local Similarity 64.3%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 KCCK--KCCK--CKC 11  
| | | | | | | | | |  
DB 13 CTCADSCCKCKCKC 26

RESULT 9  
US-08-767-026-7  
; Sequence 7, Application US/08767026  
; Patent No. 5856452  
; GENERAL INFORMATION:  
; APPLICANT: Moloney, Maurice  
; APPLICANT: Boothe, Joseph  
; APPLICANT: van Rooijen, Gijb  
; TITLE OF INVENTION: Oil Bodies and Associated Proteins as  
; TITLE OF INVENTION: Affinity Matrices  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERESKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767,026  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 9369-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-767-026-7

Query Match 58.9%; Score 46.5; DB 2; Length 254;  
Best Local Similarity 64.3%; Pred. No. 7.8;  
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 KCCK--KCCK--CKC 11  
| | | | | | | | | |  
DB 206 CTCAGSCKCKCKC 219

RESULT 10  
5240848-11  
; Patent No. 5240848

APPLICANT: KECK, PAMELA J.; CONNOLLY, DANIEL T.; FEDER, JOSEPH  
TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN VASCULAR  
PERMEABILITY FACTOR HAVING 189 AMINO ACIDS  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/337,037  
FILING DATE: 10-JUL-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 274,061  
FILING DATE: 21-NOV-1988  
SEQ ID NO: 11:  
LENGTH: 214  
5240848-11

Query Match 58.2%; Score 46; DB 4; Length 40;  
Best Local Similarity 54.5%; Pred. No. 2.5;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKCKCKCKCK 11  
I: I I I I I  
DB 28 CECPTCKCSC 38

RESULT 11  
US-08-448-418-106  
Sequence 106, Application US/08448418  
Patent No. 5837242  
GENERAL INFORMATION:  
APPLICANT: Holliger, Kaspar-Philipp  
APPLICANT: Griffiths, Andrew D  
APPLICANT: Hoogenboom, Hendricus RJM  
APPLICANT: Malmqvist, Magnus  
APPLICANT: Marks, James D  
APPLICANT: McGuinness, Brian T  
APPLICANT: Pope, Anthony R  
APPLICANT: Prospero, Terence D  
APPLICANT: Winter, Gregory P  
TITLE OF INVENTION: Multivalent and Multispecific Binding  
TITLE OF INVENTION: Proteins, Their Manufacture and Use  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun  
STREET: 6300 Sears Tower 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,418  
FILING DATE: 14-MAY-1996  
CLASSIFICATION: 435  
CLASSIFICATION: C12N 15/62, 15/70, C07K 1/00  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02492  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9225453.1  
FILING DATE: 04-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9300816.7  
FILING DATE: 16-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93303614.7  
FILING DATE: 10-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9319969.3  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:

NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32651  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-448-418-106

Query Match 55.7%; Score 44; DB 2; Length 21;  
Best Local Similarity 75.0%; Pred. No. 2.7;  
Matches 9; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 CKCK-CKC-KCK 10  
I I I I I I I I I I  
DB 6 CCKCKCKTCK 17

RESULT 12  
US-08-322-962-12  
Sequence 12, Application US/08322962  
Patent No. 5466785  
GENERAL INFORMATION:  
APPLICANT: DeFramond, Annick J  
TITLE OF INVENTION: Tissue-Preferential Promoters  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/322,962  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/071,209  
FILING DATE: 02-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/508,207  
FILING DATE: 12-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: S-18039/CGC 1479/CONT2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8615  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-322-962-12

Query Match 55.1%; Score 43.5; DB 1; Length 14;  
Best Local Similarity 57.1%; Pred. No. 2.3;  
Matches 8; Conservative 1; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11

3

Search completed: September 7, 1999, 22:38:36  
Job time: 7920 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:14 ; Search time 147.16 Seconds  
(without alignments)  
1.771 Million cell updates/sec

Title: US-09-124-280A-13  
Perfect score: 79  
Sequence: 1 CKCKCKCKCK 11

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	64	81.0	257	1 W05502	HCVM Toledo strain
2	46.5	58.9	60	1 R40209	Sequence of human
3	46.5	58.9	254	1 W64209	Olesein-metallothi
4	46.5	58.9	61	1 W61601	Human metallothion
5	45	57.0	44	1 R98208	Nucleotide used in
6	45	57.0	79	1 W63348	Secreted protein o
7	45	57.0	64	1 Y00320	Human secreted pro
8	44	55.7	244	1 P50304	Storage-specific,
9	43.5	55.1	215	1 R05102	Human vascular per
10	43.5	55.1	190	1 R08120	Mammalian glioma-d
11	43.5	55.1	190	1 R08001	Bovine vascular en
12	43.5	55.1	191	1 R08002	Human vascular end
13	43.5	55.1	164	1 R10911	Bovine vascular en
14	43.5	55.1	190	1 R22347	Rat Vascular Endot
15	43.5	55.1	214	1 R22351	Alternative form o
16	43.5	55.1	190	1 R27350	Sequence of vascul
17	43.5	55.1	214	1 R27355	Sequence of vascul
18	43.5	55.1	190	1 R27351	Sequence of vascul
19	43.5	55.1	190	1 R27352	Sequence of vascul
20	43.5	55.1	164	1 R38920	Bovine VEGF-164. I
21	43.5	55.1	165	1 R38921	Human VEGF-165. Is
22	43.5	55.1	191	1 R91076	Human vascular end
23	43.5	55.1	215	1 R91077	Human vascular end
24	43.5	55.1	232	1 R94078	Human vascular end
25	43.5	55.1	428	1 R94072	SAP(Gly4Ser)VEGF16
26	43.5	55.1	598	1 R94074	SAP(Gly4Ser)VEGF16
27	43.5	55.1	191	1 R94003	VEGF165. Vascular
28	43.5	55.1	215	1 R94002	VEGF189. Vascular
29	43.5	55.1	232	1 R94004	VEGF206. Vascular
30	43.5	55.1	192	1 R94039	VEGF165 Cys+4. Vas
31	43.5	55.1	192	1 R94040	VEGF165 Cys+2. Vas
32	43.5	55.1	461	1 W00583	VEGF165-AlaMet-SAP
33	43.5	55.1	421	1 W00584	SAP-AlaMet-VEGF165
34	43.5	55.1	443	1 W00585	SAP(Gly4Ser)4VEGF1
35	43.5	55.1	594	1 W00591	SAP-AlaMet-VEGF165
36	43.5	55.1	588	1 W00592	SAP-AlaMet-VEGF165
37	43.5	55.1	600	1 W00593	SAP-AlaMet-VEGF165
38	43.5	55.1	595	1 W00595	SAP(Gly4Ser)2VEGF1
39	43.5	55.1	612	1 W00596	SAP(Gly4Ser)2VEGF1
40	43.5	55.1	191	1 W00724	Vascular endotheli
41	43.5	55.1	190	1 W03038	Mutant vascular en
42	43.5	55.1	190	1 W03039	Mutant vascular en
43	43.5	55.1	190	1 W03040	Mutant vascular en

ALIGNMENTS

RESULT 1

W05502  
ID W05502 standard; Protein; 257 AA.  
AC W05502:  
DT 15-JAN-1997 (first entry)  
DE HCMV Toledo strain ULI33 protein (clone tol.01).  
KW CMV; HCMV; vaccine; diagnosis; ULI33.  
OS Human cytomegalovirus Toledo strain.  
PN W09630387-A1.  
PD 03-OCT-1996.  
PF 26-MAR-1996; U04100.  
PR 31-MAR-1995; US-414926.  
PA (AVIR-) AVIRON.  
PI Cha T, Spaete R;  
DR WPI: 96-455265/45.  
DR N-PSDB: T41418.  
PT New isolated human cytomegalovirus nucleic acid - from Towne and  
PT Toledo strains, used to develop prods. for the diagnosis, prevention  
PT and treatment of human CMV infections  
PS Claim 5; Page 67-68; 150pp; English.  
CC Novel protein ULI33 (W05502) is the product of an open reading  
CC frame found in a novel nucleic acid (T41418) isolated from the  
CC Toledo strain of human cytomegalovirus (HCMV). ULI33 and other  
CC novel (W05503-20) and known (W05500-01) proteins of the Toledo  
CC strain, as well as new proteins (W05521-24) from HCMV Towne, can be  
CC produced in transformed host cells and used in the prodn. of  
CC subunit vaccines against HCMV. They may be surface glycoproteins  
CC that are immunogenic or responsible for tissue tropism, or may  
CC influence the immune response of an infected individual.  
SQ Sequence 257 AA;

Query Match 81.0%; Score 64; DB 1;  
Best Local Similarity 72.7%; Pred. No. 0.099; Length 257;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKCKCKCKCK 11

Db 87 CSCCKCKCKCK 97

RESULT 2

R40209  
ID R40209 standard; protein; 60 AA.  
AC R40209:  
DT 04-FEB-1994 (first entry)  
DE Sequence of human metallothionine MT-2, class I.  
KW Metallothionine; MT-2; class I.  
OS Homo sapiens.  
PN DE4212134-A.  
PD 19-AUG-1993.  
PF 10-APR-1992; 212134.  
PR 17-FEB-1992; GB-003299.  
PA (INDE-) INDENA SPA.  
PI Bombardelli E, Ponzone C, Puglisi P;  
DR WPI: 93-265710/34.  
PT Topical compsn. for protecting tissue e.g. skin - against toxic  
PT heavy metals, contg. metal-complexing protein with high cysteine  
PT content  
PS Disclosure; Page 3; 7pp; German.  
CC Class I metallothionines are characterised by a high Cys content and  
CC the absence of aromatic AAs; a molecular weight of 6000-7000;  
CC characteristic thio-metal complexes and clusters;  
CC and a high metal content.  
SQ Sequence 60 AA;

Query Match 58.9%; Score 46.5; DB 1; Length 60;  
 Best Local Similarity 64.3%; Pred. No. 4.2;  
 Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
 | | | | | | | | | |  
 Db 12 CTCAGSCKCKCKC 25

RESULT 3

W64209 ID W64209 standard; Protein; 254 AA.  
 AC W64209;  
 DT 13-OCT-1998 (first entry)  
 DE Oleosin-metallothionein fusion protein.  
 KW Oleosin; plant; separation; oil body; thrombin; biotin; lipid; leech;  
 KW carbohydrate; cell; organelle; component; virus; fusion protein; mt-11;  
 KW metallothionein.  
 OS Brassica napus.  
 OS Arabidopsis thaliana.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO9827115-A1.  
 PD 25-JUN-1998.  
 PF 03-DEC-1997; CA0951.  
 PR 16-DEC-1996; US-767026.  
 PA (SEMB-) SEMBIOSYS GENETICS INC.  
 PI Boothe J, Moloney M, Van Rooijen G;  
 DR WPI: 98-362720/31.  
 DR N-PSDB: V44280.  
 DT Separation of target molecules from samples - by contacting with oil  
 PT bodies which associate with target molecules and then separating  
 PS Example 4; Fig 8; 94pp; English.  
 CC This sequence represents an oleosin-metallothionein fusion protein which  
 CC is constructed from a fragment of the Brassica napus oleosin protein and  
 CC a fragment of the human mt-11 protein under the control of an Arabidopsis  
 CC thaliana promoter. This protein is used in a method which allows the  
 CC separation of a target molecule from a sample comprising and involving  
 CC contacting oil bodies with a sample containing the target molecule to  
 CC allow the target molecule to associate with the oil bodies, and then  
 CC separating the oil bodies associated with the target molecule from the  
 CC sample. The method can be used for separating target molecules such as  
 CC proteins (e.g. thrombin), peptides, organic molecules (e.g. biotin),  
 CC lipids, carbohydrates, nucleic acids, cells, cell organelles, cell  
 CC components, viruses, metals (e.g. cadmium), metal ions and ions.  
 SQ Sequence 254 AA;

Query Match 58.9%; Score 46.5; DB 1; Length 254;  
 Best Local Similarity 64.3%; Pred. No. 11;  
 Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
 | | | | | | | | | |  
 Db 206 CTCAGSCKCKCKC 219

RESULT 4

W61601 ID W61601 standard; Protein; 61 AA.  
 AC W61601;  
 DT 27-OCT-1998 (first entry)  
 DE Human metallothionein HMBP-1.  
 KW Human; metallothionein; HMBP-1; metal toxicity; immune disorder;  
 KW inflammatory disease; cancers.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 28  
 FT /label= Unknown  
 FT /note= "Encoded by TYC"  
 PN WO9831795-A2.  
 PD 23-JUL-1998.

PF 30-DEC-1997; U24129.  
 PR 17-JAN-1997; US-785530.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Goli SK, Hillman JL;  
 DR WPI: 98-414095/35.  
 DR N-PSDB: V45334.  
 PT Human metallothionein, HMBP-1 - used to develop products for  
 PT diagnosis, prevention and treatment of heavy metal toxicity, cancer,  
 PT inflammatory disease and immune disorders  
 PS Claim 1: Fig 1; 53pp; English.  
 CC The human metallothionein HMBP-1 (heavy metal binding protein),  
 CC polypeptides can be used to treat heavy metal toxicity, e.g. myopathy,  
 CC encephalopathy, renal nephropathy or necrosis, liver necrosis or  
 CC cirrhosis, anaemia, myocardial damage, and pneumonitis or any other  
 CC condition or disease caused by exposure to heavy metals. They can also  
 CC be used to treat immune disorders e.g. bronchial asthma, chronic  
 CC obstructive pulmonary disease, pneumonia, multiple sclerosis, rheumatoid  
 CC arthritis, inflammatory bowel disease, chronic hepatitis, cerebral  
 CC oedema, or inflammatory disease or cancers.  
 SQ Sequence 61 AA;

Query Match 58.9%; Score 46.5; DB 1; Length 61;  
 Best Local Similarity 64.3%; Pred. No. 4.3;  
 Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 CKC--KCKCK-CKC 11  
 | | | | | | | | | |  
 Db 13 CACTGSCCKCKCKC 26

RESULT 5

R98208 ID R98208 standard; Protein; 44 AA.  
 AC R98208;  
 DT 30-DEC-1996 (first entry)  
 DE Nucleotide used in production of MSH/MoMuLV chimeric sequence.  
 KW Moloney murine leukaemia virus; gp70; 4070A retrovirus; retrovirus;  
 KW 10A1 murine leukaemia virus; NZB-9-1 murine leukaemia virus;  
 KW polyporphic MX27 provirus; targetted drug delivery; gene therapy;  
 KW single chain antibody; envelope protein; ss.  
 OS Synthetic.  
 PN WO9630504-A1.  
 PD 03-OCT-1996.  
 PR 22-MAR-1996; U03908.  
 PR 24-MAR-1995; US-409648.  
 PA (GENE-) GENETIC THERAPY INC.  
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 PI Anderson W, Chiang YL, Januszeki M, Mackrell AJ;  
 PI Zhao Y;  
 DR WPI: 96-455352/45.  
 DT Cell-targetted retroviral vector particles - having envelope protein  
 PT modified with targetting polypeptide  
 PS Example 2; Page 36; 73pp; English.  
 CC Cell targetted retroviral vector particles can be used in gene  
 CC therapy to deliver a heterologous gene to a target cell for  
 CC expression of a heterologous polypeptide in that cell. The cell  
 CC targetted retroviral vector particles comprise an envelope protein  
 CC which is modified to contain a targetting polypeptide (a single chain  
 CC antibody), or in the case of moloney murine leukaemia virus  
 CC (MoMuLV), alpha melanotropin-stimulating hormone (MSH). Two  
 CC oligonucleotides (R98207, R98208) were used to substitute sequences in  
 CC MoMuLV for MSH sequences. This oligonucleotide was used to replace  
 CC residues G80-P88 of MoMuLV envelope protein (See W04248).  
 SQ Sequence 44 AA;

Query Match 57.0%; Score 45; DB 1; Length 44;  
 Best Local Similarity 54.5%; Pred. No. 5.1;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCKC 11  
 | | | | | | | | | |



Db 30 CTCCTCACCC 40

RESULT 6

W69348

AC W69348 standard; Protein; 79 AA.

DT 25-NOV-1998 (first entry)

DE Secreted protein of clone FG340\_1.

KW Secreted protein; nutritional activity; immune stimulant;

KW immune suppressor; haematopoiesis regulator; tissue growth activity;

KW activin/inhibin activity; chemotactic/chemokinetic activity;

KW haemostatic activity; thrombolytic activity; receptor/ligand activity;

KW anti-inflammatory activity; cadherin suppressor; tumour inhibitor;

KW tumour invasion suppressor; therapy.

OS Homo sapiens.

PN W09837094-A2.

PD 27-AUG-1998.

PF 24-FEB-1998: U03595.

PR 23-FEB-1998: US-028168.

PR 24-FEB-1997: US-804561.

PA (GENY ) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,

PI Racie LA, Spaulding V, Treacy M;

DR WPI: 98-467491/40.

DR N-PSDB: V58364.

PT Secreted proteins with biological activity - and encoding

PT polynucleotide(s), useful e.g. in therapy and diagnosis of medical

PT conditions and to identify protein agonists or antagonists

PS Claim 39: Page 89; 111pp; English.

CC This sequence is a secreted protein of the invention. This sequence

CC is encoded by the DNA of clone FG340\_1, which was isolated from a human

CC adult brain cDNA library. The polypeptides are predicted to have

CC useful biological activities which would make them suitable for treating,

CC preventing or ameliorating medical conditions in humans and animals,

CC although no supporting biological data is given. Suggested activities

CC include nutritional activity, immune stimulating or suppressing activity,

CC haematopoiesis regulating activity, tissue growth activity,

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition

CC activity or other activities. They may also be used for diagnostic

CC purposes. The polynucleotides are useful to produce the proteins or

CC generate probes or primers to identify and/or amplify similar genes

CC e.g. species homologues. They are also useful for gene therapy and to

CC produce transgenic animals with altered gene expression.

SQ Sequence 79 AA;

Query Match 57.0%; Score 45; DB 1; Length 79;

Best Local Similarity 54.5%; Pred. No. 7.7;

Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCK 11

Db 9 CVCVCVCVCAC 19

RESULT 7

Y00320

ID Y00320 standard; Protein; 64 AA.

AC Y00320;

DT 04-MAY-1999 (first entry)

DE Human secreted protein encoded by gene 64.

KW Human; secreted protein; fusion protein; gene therapy; protein therapy;

KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;

KW developmental abnormality; foetal deficiency; blood; allergy; renal;

KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;

KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;

KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;

KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;

KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN W09906423-A1.

PD 11-FEB-1998.

PF 29-JUL-1998; U15949.

PR 19-AUG-1997; US-056730.

PR 30-JUL-1997; US-054209.

PR 30-JUL-1997; US-054211.

PR 30-JUL-1997; US-054212.

PR 30-JUL-1997; US-054213.

PR 30-JUL-1997; US-054214.

PR 30-JUL-1997; US-054215.

PR 30-JUL-1997; US-054217.

PR 30-JUL-1997; US-054218.

PR 30-JUL-1997; US-054234.

PR 30-JUL-1997; US-054236.

PR 18-AUG-1997; US-055968.

PR 18-AUG-1997; US-055969.

PR 18-AUG-1997; US-055972.

PR 19-AUG-1997; US-056534.

PR 19-AUG-1997; US-056543.

PR 19-AUG-1997; US-056554.

PR 19-AUG-1997; US-056561.

PR 19-AUG-1997; US-056727.

PR 19-AUG-1997; US-056729.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Carter KC, Endress GA, Fan P, Feng P, Kyaw H, Lafleur DW,

PI Li Y, Moore PA, Rosen CA, Ruben SM, Shi Y, Wei Y,

PI zeng Z;

DR WPI: 99-153691/13.

DR N-PSDB: X20475.

PT New isolated human genes and the secreted polypeptides they encode -

PT useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders

PS Claim 11; Page 283; 312pp; English.

CC This sequence represents a secreted human protein encoded by the gene

CC clone detailed in the descriptor line. The gene can be used to generate

CC fusion proteins by linking to the gene to a human immunoglobulin FC

CC portion (e.g. X20403) for increasing the stability of the fused protein

CC as compared to the human protein only.

CC The invention relates to 83 novel genes and their fragments (nucleic acid

CC sequences: X20412-X20499; amino acid sequences Y00258-Y00377) which

CC are useful for preventing, treating or ameliorating medical conditions

CC e.g. by protein or gene therapy. Also, pathological conditions can be

CC diagnosed by determining the amount of the new polypeptides in a sample

CC or by determining the presence of mutations in the new polynucleotides.

CC Specific uses are described for each of the 83 polynucleotides, based on

CC which tissues they are most highly expressed in (see X20412 for described

CC uses).

SQ Sequence 64 AA;

Query Match 57.0%; Score 45; DB 1; Length 64;

Best Local Similarity 54.5%; Pred. No. 6.6;

Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CKCKCKCKCK 11

Db 49 CVCVCVCVCVC 59

RESULT 8

P50304

ID P50304 standard; Protein; 244 AA.

AC P50304;

DT 22-OCT-1991 (first entry)

DE Storage-specific, late schizont merozoite malaria antigen.

KW Malaria; vaccine.

OS Plasmodium spp.

PN W08503725-A.

PD 29-AUG-1985.

PF 20-FEB-1985; G00072.

PR 20-FEB-1984; GB-004378.

PA (BIOJ ) BIOGEN NV.

PI Mach B, Perrin L, McGarvey M, Cheung A, Shaw A;

DR WPI: 85-223371/36.  
 DR N-PSDB; N50355.  
 PT Prodn. of antigens of Plasmodium species, esp. of falciparum - by  
 PT recombinant DNA methods giving polypeptide(s) for protecting  
 PT against malaria or for diagnosis.  
 PS Disclosure; Fig 6; 49pp; English.  
 CC The sequence encodes a Plasmodium falciparum, Plasmodium vivax,  
 CC Plasmodium malariae and Plasmodium ovale antigen which may be used  
 CC in the diagnosis of malaria and as a vaccine against malaria.  
 SQ Sequence 244 AA;

Query Match 55.7%; Score 44; DB 1; Length 244;  
 Best Local Similarity 70.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 KCCKCKCKC 11  
 | |||||  
 Db 33 KAMCKCKNC 42

RESULT 9  
 R05102  
 ID R05102 standard; protein; 215 AA.  
 AC R05102;  
 DT 05-OCT-1990 (first entry)  
 DE Human vascular permeability factor  
 KW human vascular permeability factor (hVPF); zinc-finger;  
 KW human histiocytic lymphoma cell line U-937; ss.  
 OS synthetic.  
 FH Key  
 FT Location/Qualifiers  
 FT peptide 1..26  
 FT /label=signal peptide  
 FT protein 27..215  
 FT /label=hVPF  
 FT modified\_site 101..101  
 FT /label=N-glycosylated Asn  
 FT domain 83..116  
 FT /label=Putative Zn-finger domain

EP-370989-A.  
 PD 30-MAY-1990.  
 PF 20-NOV-1989; 870181.  
 PR 21-NOV-1988; US-274061.  
 PA (MONS) Monsanto Co.  
 PI Keck PJ, Feder J, Connolly DT, Olander JV;  
 DR WPI: 90-165658/22.  
 DR N-PSDB: Q04474.  
 PT Human vascular permeability factor -  
 PT obtd. by culture of human histiocytic lymphoma cell line U-937  
 PT and by recombinant DNA techniques  
 PS Claim 5; Page 19; 24pp; English.  
 CC Sequence includes 4 amino acid segments which correspond to a U-937  
 CC secreted protein all ready sequenced (see US-274061). These are the  
 CC following: amino acids 1-10, 17-23, 33-45 and 46-56.  
 SQ Sequence 215 AA;

Query Match 55.1%; Score 43.5; DB 1; Length 215;  
 Best Local Similarity 34.6%; Pred. No. 23;  
 Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CK 11  
 ||| || || ||  
 Db 185 CKCKNTDSCKARQLELNERTC 210

RESULT 10  
 R08120  
 ID R08120 standard; protein; 190 AA.  
 AC R08120;  
 DT 24-FEB-1991 (first entry)  
 DE Mammalian glioma-derived growth factor (GDGF) gene product.  
 KW Mitogenesis; wound healing.

OS Rattus rattus.  
 PN EP-399816-A.  
 PD 28-NOV-1990.  
 PF 23-MAY-1990; 305637.  
 PR 24-MAY-1989; US-356477.  
 PR 30-MAR-1990; US-500421.  
 PA (MERI ) MERCK & CO INC.  
 PI Bayne ML, Conn GL, Thomas KA;  
 DR WPI: 90-356428/48.  
 DR N-PSDB: Q06741.  
 PT Glioma-derived growth factor - stimulates mitogenesis of  
 PT mammalian endothelial cells and is used in vitro or in vivo to  
 PT stimulate growth  
 PS Claim 19; Fig 1; 19pp; English.  
 CC GDGF stimulates mitogenesis and vascular endothelial cell growth,  
 CC useful in promoting wound healing or tissue growth. It allows vascular  
 CC explants to be grown on a GDGF treated surface for implantation in a  
 CC host. The sequence was isolated from the GS-9L rat cells.  
 SQ Sequence 190 AA;

Query Match 55.1%; Score 43.5; DB 1; Length 190;  
 Best Local Similarity 34.6%; Pred. No. 21;  
 Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CK 11  
 ||| || || ||  
 Db 160 CKCKNTDSCKARQLELNERTC 185

RESULT 11  
 R08001  
 ID R08001 standard; protein; 190 AA.  
 AC R08001;  
 DT 27-FEB-1991 (first entry)  
 DE Bovine vascular endothelial growth factor as deduced from clone  
 DE lambda.vgef.6.  
 KW VEGF; pituitary follicular cell.  
 OS Bos taurus.  
 FH Key  
 FT Location/Qualifiers  
 FT peptide 1..26  
 FT /label=signal peptide  
 FT protein 27..191  
 FT /label=bVEGF  
 PN W09013649-A.  
 PD 15-NOV-1990.  
 PF 09-MAY-1990; U02585.  
 PR 12-MAY-1989; US-351117.  
 PR 21-JUN-1989; US-369424.  
 PR 04-AUG-1989; US-389722.  
 PA (GETH ) GENENTECH INC.  
 PI Ferrara N, Leung DWH;  
 DR WPI: 90-361484/48.  
 DR N-PSDB: Q06600.  
 PT DNA encoding vascular endothelial cell growth factor - having  
 PT selective action for treatment of vascular conditions in absence  
 PT of excessive tissue proliferation.  
 PS Claim 37; Fig 2; 51pp; English.  
 CC A probe (Q06599) was used to screen a bovine pituitary follicular  
 CC cell library for cDNA clones encoding the growth factor. One  
 CC positive clone, designated lambda.vgef.6 was identified. This  
 CC clone was completely sequenced (Q06600) after subcloning into the  
 CC PRK5 vector. This AA sequence was deduced from that sequence. The  
 CC expression vector, PRK5.vgef.6, was used to produce recombinant  
 CC VEGF for use in the treatment of conditions in which a selective  
 CC action on the vascular endothelial cells in the absence of excess-  
 CC ive tissue proliferation is desirable.  
 CC See also R08002.  
 SQ Sequence 190 AA;

Query Match 55.1%; Score 43.5; DB 1; Length 190;  
 Best Local Similarity 34.6%; Pred. No. 21;

Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CK 11  
 DB 160 CKCSCKNTDSRCKARQLELNERTC 185

## RESULT 12

ID R08002 standard; protein; 191 AA.  
 AC R08002;  
 DT 27-FEB-1991 (first entry)  
 DE Human vascular endothelial growth factor as deduced from clone  
 DE lambda.veg.21.  
 KW VEGF; leukaemia.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..26  
 FT /label= signal peptide  
 FT protein 27..191  
 FT /label= hVEGF  
 FN W09013649-A.  
 PD 15-NOV-1990.  
 PF 09-MAY-1990; U02585.  
 PR 12-MAY-1989; US-351117.  
 PR 21-JUN-1989; US-369424.  
 PR 04-AUG-1989; US-389722.  
 PA (GETH ) GENENTECH INC.  
 PI Ferrara N, Leung DW;  
 DR WPI: 90-361484/48.  
 DR N-PSDB; Q07006.  
 PT DNA encoding vascular endothelial cell growth factor - having  
 PT selective action for treatment of vascular conditions in absence  
 PT of excessive tissue proliferation.  
 PS Claim 38; Fig 10; 51pp; English.  
 CC Clone lambda.veg.6 (Q06600), representing the full length DNA of  
 CC bovine VEGF was used to screen cDNA library pred. from RNA ex-  
 CC tracted from human leukemia cell line HL60 (ATCC CCL240) for cDNA  
 CC clones encoding the growth factor. Five positive clones were  
 CC identified, one of which was designated lambda.veg.216. This  
 CC clone was completely sequenced (Q07006). This AA sequence was  
 CC deduced from that sequence. The expression vector, p.veg.21, was  
 CC used to produce recombinant VEGF for use in the treatment of  
 CC conditions in which a selective action on the vascular endothelial  
 CC cells in the absence of excessive tissue proliferation is desirable.  
 CC See also R08001.  
 SQ Sequence 191 AA;

Query Match 55.1%; Score 43.5; DB 1; Length 191;  
 Best Local Similarity 34.6%; Pred. No. 21;  
 Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CK 11  
 DB 161 CKCSCKNTDSRCKARQLELNERTC 186

## RESULT 13

ID R10911 standard; Protein; 164 AA.  
 AC R10911;  
 DT 08-MAY-1991 (first entry)  
 DE Bovine vascular endothelial cell growth factor 164.  
 DE Bovine vascular endothelial cell growth factor; angiogenesis;  
 KW wound healing; bVEGF; PDGF.  
 OS Bos taurus.  
 PN W09102058-A.  
 PD 21-FEB-1991.  
 PF 27-JUL-1990; U04227.  
 PR 27-JUL-1989; US-387545.  
 PR 14-DEC-1989; US-450883.  
 PA (CALB-) CALIF BIOTECHN INC.

PI Tischer ER, Abrahamam, Fiddes JC, Mitchell RL;  
 DR WPI: 91-073534/10.  
 DR N-PSDB; Q10791.

PT DNA encoding vascular endothelial cell growth factor - used for  
 PT producing the factor for angiogenesis and re-endothelialisation  
 PT in wound healing  
 PS Disclosure; Fig 6(1-3); 94pp; English.  
 CC Bovine folliculo stellate cells were used in the process of  
 CC obtaining cDNA encoding bVEGF (164 amino acids from). The probes  
 CC represented in Q10806 and -07 were used in the screening procedures.  
 CC See Q10796 for bVEGF120 which is obtained by alternative splicing of  
 CC this sequence, i.e. bases 342-473 (amino acids 115-158) are spliced.  
 CC The product can be used for angiogenesis and re-endothelialisation  
 CC of inner vascular surfaces in wound healing, e.g. treatment of full-  
 CC thickness wounds such as dermal ulcers, venous ulcers and diabetic  
 CC ulcers, burns, in surgery, in balloon angioplasty and for the in  
 CC vitro culturing of endothelial cells. Hybrid growth factors of PDGF  
 CC and VEGF can exhibit a mitogenic profile between each factor and  
 CC can be used for wound healing or as inhibitors of angiogenesis for  
 CC e.g. preventing the growth of tumours.  
 CC VEGF analogues in which CYS residues are substd. are more stable.  
 CC See also Q10791-93; Q10796-97; Q10806-08 and Q11099.  
 SQ Sequence 164 AA;

Query Match 55.1%; Score 43.5; DB 1; Length 164;  
 Best Local Similarity 34.6%; Pred. No. 19;  
 Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CK 11  
 DB 134 CKCSCKNTDSRCKARQLELNERTC 159

## RESULT 14

ID R22347 standard; Protein; 190 AA.  
 AC R22347;  
 DT 29-JUL-1992 (first entry)  
 DE Rat Vascular Endothelial Growth Factor A-subunit.  
 KW VEGF-I; mammalian glioma cell; conditioned medium; heterodimer;  
 KW homodimer; mitogenesis; VEGF-II; vascular repair;  
 KW blood vessel implant.  
 OS Rattus.  
 FH Key Location/Qualifiers  
 FT peptide 1..26  
 FT /label= signal  
 FT protein 27..190  
 FT /label= mature\_VEGF\_A\_monomer  
 FT modified\_site 100..102  
 FT /label= N-glycosylation  
 FT peptide 27..41  
 FT /label= T27  
 FT /note= "tryptic peptide"  
 FT peptide 42..48  
 FT /label= T41  
 FT /note= "tryptic peptide"  
 FT peptide 49..81  
 FT /label= T65  
 FT /note= "tryptic peptide"  
 FT peptide 127..135  
 FT /label= T18  
 FT /note= "tryptic peptide"  
 FT peptide 127..132  
 FT /label= T15  
 FT /note= "tryptic peptide"  
 FT peptide 136..150  
 FT /label= T22  
 FT /note= "tryptic peptide"  
 FT peptide 151..170  
 FT /label= T38  
 FT /note= "tryptic peptide"  
 FT peptide 175..184

```
FT peptide /label= T32
FT /note= "tryptic peptide"
FT 27. .31
FT /label= V11A
FT /note= "v8 Protease cleavage peptide"
FT 39. .62
FT /label= V40
FT /note= "v8 Protease cleavage peptide"
FT 68. .98
FT /label= V30
FT /note= "v8 Protease cleavage peptide"
FT 99. .118
FT /label= V18A
FT /note= "v8 Protease cleavage peptide"
FT 119. .139
FT /label= V18B
FT /note= "v8 Protease cleavage peptide"
FT 148. .177
FT /label= V21
FT /note= "v8 Protease cleavage peptide"
FT 178. .190
FT /label= V11
FT /note= "v8 Protease cleavage peptide"
FT 36. .41
FT /label= L12
FT /note= "Leu C cleavage peptide"
FT 42. .133
FT /label= L42
FT /note= "Leu C cleavage peptide"
FT 134. .150
FT /label= L16
FT /note= "Leu C cleavage peptide"
FT 173. .190
FT /label= L20
FT /note= "Leu C cleavage peptide"
FT 27. .43
FT /label= CB26
FT /note= "Cyanogen Bromide cleavage peptide"
FT 107. .119
FT /label= CB18-19
FT /note= "Cyanogen Bromide cleavage peptide"
FT EP-476983-A.
FT 25-MAR-1992.
FT 18-SEP-1991; 308489.
FT 21-SEP-1990; US-586640.
FT 21-SEP-1990; US-586638.
FT (MERI ) MERCK & CO INC.
FT Bayne ML, Conn GL, Thomas KA;
FT WPI: 92-098641/13.
FT N-PSDB; Q23038.
FT Vascular endothelial cell growth factor II - used as coating for
FT artificial blood vessels or to promote tissue repair
FT Disclosure; Fig 4; 38pp; English.
FT The VEGF-I A-subunit was found to be identical to the VEGF-II
FT A-subunit. The isolated A-monomer was treated with trypsin, CNBr,
FT Lys C or V8 protease. The peptides resulting from the digestions
FT were isolated and sequenced using the Edman degradation reaction in
FT a gas phase sequenator. In addition to the peptide fragments which
FT were directly sequenced, the full-length protein sequence was
FT deduced from the cDNA encoding it.
FT See also Q23040-Q23059.
FT Sequence 190 AA;
SQ
```

```
Query Match 55.1%; Score 43.5; DB 1; Length 190;
Best Local Similarity 34.6%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CKC 11
   ||| || ||
Db 160 CKCSCKNTDSRCRKAROLELNERTC 185
```

```
RESULT 15
R22351
ID R22351 standard; Protein; 214 AA.
AC R22351;
DT 29-JUL-1992 (first entry)
DE Alternative form of VEGF mature A-subunit with 188 amino acids.
KW Rat glioma cell; GS-9L; conditioned medium; heterodimer; VEGF-II;
KW homodimer; mitogenesis; vascular repair; blood vessel implant;
KW polymerase chain reaction; alternative splicing.
OS Rattus.
FH Key Location/Qualifiers
FT peptide 1..26
FT /label= signal
FT protein 27..214
FT /label= VEGF_A-subunit
FT /note= "188 amino acids long"
PN EP-476983-A.
PD 25-MAR-1992.
PF 18-SEP-1991; 308489.
PR 21-SEP-1990; US-586640.
PR 21-SEP-1990; US-586638.
PA (MERI ) MERCK & CO INC.
PI Bayne ML, Conn GL, Thomas KA;
DR WPI: 92-098641/13.
DR N-PSDB; Q23051.
PT Vascular endothelial cell growth factor II - used as coating for
PT artificial blood vessels or to promote tissue repair
PT Example 9; Page 14 and Fig 4; 38pp; English.
CC Multiple cDNAs encoding alternative forms of the VEGF A-subunit
CC were amplified using PCR primers as in Q23049 and Q23050. Three
CC sets of clones were identified. Clone #12 encoded the 164 amino acid
CC secreted form of VEGF A-subunit (see R22347). Clone #14 has a 135 bp
CC deletion and thus encodes a 120 amino acid form (see R22348) and
CC Clone #16 has a 72bp insertion and encodes a 188 amino acid mature
CC protein (R22351). The 188 amino acid secreted form of the VEGF
CC A-subunit has Asn140 converted to Lys140.
CC See also Q23038-Q23050 and Q23052-Q23059.
SQ Sequence 214 AA;
```

```
Query Match 55.1%; Score 43.5; DB 1; Length 214;
Best Local Similarity 34.6%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 1; Indels 15; Gaps 2;

QY 1 CKCKCK-----CK-----CKC 11
   ||| || ||
Db 184 CKCSCKNTDSRCRKAROLELNERTC 209
```

Search completed: September 7, 1999, 20:37:15  
Job time: 18464 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:15 ; Search time 147.16 Seconds  
(without alignments)  
0.966 Million cell updates/sec

Title: US-09-124-280A-14

Perfect score: 39

Sequence: 1 CKFKKC 6

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	6	1 R71785	Peptide neutralisi
2	39	100.0	6	1 W21602	Antibiotic potenti
3	36	92.3	72	1 P91995	Part of v-erb A on
4	36	92.3	68	1 R08246	GTG32 chimeric glu
5	36	92.3	68	1 R08248	GTG33 chimeric glu
6	36	92.3	636	1 R07883	gag protein precu
7	36	92.3	410	1 P80936	Sequence encode by
8	36	92.3	490	1 P80926	Sequence of the hu
9	36	92.3	345	1 P80922	Sequence of the ca
10	36	92.3	456	1 P80921	Sequence encoded b
11	36	92.3	490	1 R26899	Herba-T sequence.
12	36	92.3	514	1 R47821	Rat thyroid hormon
13	36	92.3	579	1 R71565	EAR-1r gene produc
14	36	92.3	410	1 R78318	Human thyroid horm
15	36	92.3	410	1 W41493	LambdaMe2 containi
16	32	82.1	439	1 R33745	XR4. DNA encoding
17	32	82.1	441	1 R41875	Steroid hormone re
18	32	82.1	598	1 R48831	Sequence of nuclea
19	32	82.1	441	1 R89214	Peroxisome prolife
20	32	82.1	475	1 R92478	Peroxisome prolife
21	32	82.1	440	1 R92479	Peroxisome prolife
22	32	82.1	628	1 R92057	Apoptotic cerebral
23	32	82.1	494	1 R93323	Peroxisome prolife
24	32	82.1	477	1 R99324	Peroxisome prolife
25	32	82.1	475	1 R99325	Peroxisome prolife
26	32	82.1	505	1 R99326	Peroxisome prolife
27	32	82.1	477	1 R99327	Peroxisome prolife
28	32	82.1	475	1 R99328	Peroxisome prolife
29	32	82.1	626	1 W16398	Human neuron-deriv
30	32	82.1	225	1 W08079	Human fibroblast g
31	32	82.1	225	1 W29287	Human fibroblast g
32	32	82.1	225	1 W53012	Fibroblast growth
33	32	82.1	225	1 W53027	Mouse fibroblast g
34	32	82.1	9	1 W60435	Tumour homing pept
35	32	82.1	475	1 W99596	Mouse peroxisome p
36	32	82.1	440	1 W99597	Mouse peroxisome p
37	32	82.1	9	1 W93762	Human Kaposi's sar
38	31	79.5	1100	1 R15783	B.thuringiensis to
39	31	79.5	956	1 R15784	B.thuringiensis to
40	31	79.5	914	1 R15785	B.thuringiensis to
41	31	79.5	9	1 W60434	Tumour homing pept
42	31	79.5	511	1 W78475	Autographa califor
43	31	79.5	530	1 W78476	Baculovirus ISP pr

ALIGNMENTS

RESULT 1

ID R71785 R71785 standard; peptide; 6 AA.  
AC R71785;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide bond 1..6  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WFI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
PS Claim 16; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock CC and also in vitro to detoxify vaccines, drug solutions, injectable CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new CC peptides. 6 AA;  
SQ Sequence 6 AA;

Query Match 100.0%; Score 39; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
| | | | |  
Db 1 CKFKKC 6

RESULT 2

ID W21602 W21602 standard; peptide; 6 AA.  
AC W21602;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #14.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
OS permeability; outer bacterial membrane.  
FH Key Location/Qualifiers  
FT disulfide bond 1..6  
PN W09638163-A.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WFI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
PS Claim 18; Page 25; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 6 AA;

Query Match 100.0%; Score 39; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 1 CKFKKC 6

## RESULT 3

P91995  
 ID P91995 standard; protein; 72 BP.

AC P91995;  
 DT 05-MAR-1990 (first entry)  
 DE Part of v-erb A oncogene product.  
 KW V-erb A oncogene; cysteine-rich DNA binding domain;  
 KW avian erythroblastosis virus.  
 OS Avian erythroblastosis virus.  
 PN W08909223-A.

PD 05-OCT-1989.  
 PF 24-MAR-1989; U01238.  
 PR 30-MAR-1988; US-176107.

PA (ARCH-) Arch Development Corp.

PI Liao S, Chang C;

DR WPI; 89-309501/42.

PT New DNA encoding new androgen receptor and TR2 polypeptide(s) - able  
 PT to bind DNA, and derived antibodies, useful for receptor assay and  
 PT purification.

PS Disclosure; fig. 5: 60pp; English.

CC The sequence is part of the v-erb oncogene product and is homologous  
 CC with the cysteine-rich DNA binding domain of human androgen receptor,  
 CC glucocorticoid receptor, mineralocorticoid receptor, progesterone  
 CC receptor, TR2, rat AR, and chick vitamin D receptor.  
 CC See also P91991 - P91996.

SQ Sequence 72 AA;

Query Match 92.3%; Score 36; DB 1; Length 72;  
 Best Local Similarity 83.3%; Pred. No. 8.7;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 58 CRFKKC 63

## RESULT 4

R08246  
 ID R08246 standard; protein; 68 AA.

AC R08246;  
 DT 01-FEB-1991 (first entry)  
 DE GTC32 chimeric glucocorticoid/thyroid receptor.  
 KW Steroid; thyroid; hormone response element; HER;  
 KW ligand identification; chimera; orphan receptor protein; ORP.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT domain 1..21  
 FT /label=finger\_1  
 FT region 22..38  
 FT /label=linker  
 FT domain 39..63

FT W09011273-A.  
 PN /label=finger\_2

PD 04-OCT-1990.

PF 16-MAR-1990; U01428.

PR 17-MAR-1989; US-325240.

PA (SALK ) SALK INST FOR BIOL STUD.

PI Evans RM, Kazuhiko U;

DR WPI; 90-320209/42.

PT Steroid and thyroid hormone response elements - used to develop  
 PT assays for identifying ligands for orphan hormone receptors

PS Disclosure; Fig 2; 48pp; English.

CC To identify amino acids that allow the receptors to discriminate  
 CC between their cognate HREs, a variety of chimeric DNA binding  
 CC domains were constructed. Portions of the loops and linkers were  
 CC exchanged between the hGR and hGRbeta or human retinoic acid  
 CC receptor alpha (hRARalpha). GTC32 has Loop 1 switched. The mutant  
 CC is active (29%-TREP) upon induction of a luciferase activity. Since  
 CC the specificity has not changed, these affected amino acids are  
 CC apparently not critical in discriminating between GREs and TREs or  
 CC TREP from GREs, but rather are involved in common aspects of HRE  
 CC recognition. Nine amino acids in two clusters were not changed and  
 CC may therefore impart specificity. These are the proximal (P) and  
 CC distal (D) region.

CC See also R08245-51, R07366-86, R07391, -95 and -96.

SQ Sequence 68 AA;

Query Match 92.3%; Score 36; DB 1; Length 68;  
 Best Local Similarity 83.3%; Pred. No. 8.3;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 58 CRFKKC 63

## RESULT 5

R08248  
 ID R08248 standard; protein; 68 AA.

AC R08248; 1991 (first entry)

DE GTC33 chimeric glucocorticoid/thyroid receptor.

KW Steroid; thyroid; hormone response element; HER;

KW ligand identification; chimera; orphan receptor protein; ORP.

OS Homo sapiens.

FH Key Location/Qualifiers

FT domain 1..21  
 FT /label=finger\_1  
 FT region 22..38  
 FT /label=linker  
 FT domain 39..63  
 FT /label=finger\_2

PN W09011273-A.

PD 04-OCT-1990.

PF 16-MAR-1990; U01428.

PR 17-MAR-1989; US-325240.

PA (SALK ) SALK INST FOR BIOL STUD.

PI Evans RM, Kazuhiko U;

DR WPI; 90-320209/42.

PT Steroid and thyroid hormone response elements - used to develop  
 PT assays for identifying ligands for orphan hormone receptors

PS Disclosure; Fig 2; 48pp; English.

CC To identify amino acids that allow the receptors to discriminate  
 CC between their cognate HREs, a variety of chimeric DNA binding  
 CC domains were constructed. Portions of the loops and linkers were  
 CC exchanged between the hGR and hGRbeta or human retinoic acid  
 CC receptor alpha (hRARalpha). GTC33 has Loop 2 switched. The mutant  
 CC is active (61%-TREP) upon induction of a luciferase activity. Since  
 CC the specificity has not changed, these affected amino acids are  
 CC apparently not critical in discriminating between GREs and TREs or  
 CC TREP from GREs, but rather are involved in common aspects of HRE  
 CC recognition. Nine amino acids in two clusters were not changed and  
 CC may therefore impart specificity. These are the proximal (P) and

CC distal (D) region.  
CC See also R08245-56, R07366-86, R07391, -95 and -96.  
SQ Sequence 68 AA;

Query Match 92.3%; Score 36; DB 1; Length 68;  
Best Local Similarity 83.3%; Pred. No. 8.3;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 58 CRFKKC 63

RESULT 6  
R07683  
ID R07683 standard; protein; 636 AA.  
AC R07683;  
DE 18-FEB-1991 (first entry)  
DE gag protein precursor, p75gag-v-erba.  
KW Retroviral protease activity assay; gag protein precursor;  
KW p-75gag-v-erba; bovine leukaemia virus.  
OS Avian sp.  
PN J02249498-A.  
PD 05-OCT-1990.  
PF 22-MAR-1989; JP-066772.  
PR 22-MAR-1989;  
PA (TOYJ ) TOSOH CORP.  
PA (NISK ) NIPPON SEIBUTSU KAG KK.  
DR WPI: 90-345010/46.  
PT Assay method of retro-virus protease activity - using label  
PT binding avian erythroblastosis virus derived gag protein  
PT precursor p75-gag-v-erba as substrate  
PS Disclosure: fig 3; 6pp; Japanese.  
CC This protein , which is an avian erythroblastosis-derived gag  
CC protein precursors, is used in an assay of retroviral protease  
CC activity. It is contacted with a virus, e.g. AMV, BLV, EIAV,  
CC HTLV-1 or HIV, and the resultant reaction is measured. This  
CC assay has a potential use in the development of a protease in-  
CC hibitor or antiviral drug.  
SQ Sequence 636 AA;

Query Match 92.3%; Score 36; DB 1; Length 636;  
Best Local Similarity 83.3%; Pred. No. 48;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 348 CRFKKC 353

RESULT 7  
P80936  
ID P80936 standard; protein; 410 AA.  
AC P80936;  
DE 29-DEC-1990 (first entry)  
DE Sequence encode by thyroid hormone receptor cDNA from rat brain clone  
DE rbeA12  
KW Hormone receptor; hormone-binding; transcription activation.  
OS Rat.  
PN WO8803168-A.  
PD 05-MAY-1988.  
PF 23-OCT-1987; U02782.  
PR 20-OCT-1987; US-108471.  
PA (SALK) Salk Inst for Biol Stud.  
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;  
DR WPI: 88-133242/19.  
DR N-PSDB; N80923.  
PT Recombinant DNA encoding hormone receptors -  
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone  
PT and novel hormone receptors  
PS Example; Fig VII-1(B); 243pp; English.

CC DNAs encoding hormone receptors and the hormone receptors themselves are  
CC claimed. The DNA can be used to make the hormone receptor proteins and  
CC functional modified forms in quantities not previously possible. The  
CC receptor proteins can be used to screen cpds. for receptor-agonist or  
CC receptor-antagonist activity. They can also be used in diagnostic assays.  
SQ Sequence 410 AA;

Query Match 92.3%; Score 36; DB 1; Length 410;  
Best Local Similarity 83.3%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 110 CRFKKC 115

RESULT 8  
P80926  
ID P80926 standard; protein; 490 AA.  
AC P80926;  
DE 29-DEC-1990 (first entry)  
DE Sequence of the human thyroid receptor HERBA 8.7  
KW Hormone receptor; hormone-binding; transcription activation.  
OS Homo sapiens.  
PN WO8803168-A.  
PD 05-MAY-1988.  
PF 23-OCT-1987; U02782.  
PR 20-OCT-1987; US-108471.  
PA (SALK) Salk Inst for Biol Stud.  
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;  
DR WPI: 88-133242/19.  
DR N-PSDB; N80919.  
PT Recombinant DNA encoding hormone receptors -  
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone  
PT and novel hormone receptors  
PS Claim 17; Fig III-7; 243pp; English.  
CC DNAs encoding hormone receptors and the hormone receptors themselves are  
CC claimed. The DNA can be used to make the hormone receptor proteins and  
CC functional modified forms in quantities not previously possible. The  
CC receptor proteins can be used to screen cpds. for receptor-agonist or  
CC receptor-antagonist activity. They can also be used in diagnostic assays.  
CC The sequence of thyroid receptor hPA is related to HERBA 8.7.  
SQ Sequence 490 AA;

Query Match 92.3%; Score 36; DB 1; Length 490;  
Best Local Similarity 83.3%; Pred. No. 39;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 110 CRFKKC 115

RESULT 9  
P80922  
ID P80922 standard; protein; 345 AA.  
AC P80922;  
DE 29-DEC-1990 (first entry)  
DE Sequence of the carboxy-terminal of the v-erba oncogene product  
KW Hormone receptor; hormone-binding; transcription activation;  
KW thyroid receptor.  
OS Virus.  
PN WO8803168-A.  
PD 05-MAY-1988.  
PF 23-OCT-1987; U02782.  
PR 20-OCT-1987; US-108471.  
PA (SALK) Salk Inst for Biol Stud.  
PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;  
DR WPI: 88-133242/19.  
PT Recombinant DNA encoding hormone receptors -  
PT comprising glucocorticoid, mineralocorticoid, thyroid hormone

PT and novel hormone receptors  
 PS Example; Fig III-2; 243pp; English.  
 CC DNAs encoding hormone receptors and the hormone receptors themselves are  
 CC claimed. The DNA can be used to make the hormone receptor proteins and  
 CC functional modified forms in quantities not previously possible. The  
 CC receptor proteins can be used to screen cpds. for receptor-agonist or  
 CC receptor-antagonist activity. They can also be used in diagnostic assays.  
 SQ Sequence 345 AA;

Query Match 92.3%; Score 36; DB 1; Length 345;  
 Best Local Similarity 83.3%; Pred. NO. 30;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 59 CRFKKC 64

## RESULT 10

ID R26899 standard; protein; 456 AA.  
 AC R26899; 1993 (first entry)  
 DT 29-DEC-1990 (first entry)  
 DE Sequence encoded by human placenta thyroid receptor c-erbA cDNA  
 KW Hormone receptor; hormone-binding; transcription activation.  
 OS Homo sapiens.  
 PN W08803168-A.  
 PD 05-MAY-1988.  
 PF 23-OCT-1987; U02782.  
 PR 20-OCT-1987; US-108471.  
 PA (SALK) Salk Inst for Biol Stud.  
 PI Evans RM, Weinberger CA, Hollenberg SM, Giguere V;  
 DR WPI: 88-133242/19.  
 DR N-PSDB: R80918.  
 PT Recombinant DNA encoding hormone receptors -  
 PT comprising glucocorticoid, mineralocorticoid, thyroid hormone  
 PT and novel hormone receptors  
 PS Claim 17; Fig III-1(B)-1 and -2; 243pp; English.  
 CC DNAs encoding hormone receptors and the hormone receptors themselves are  
 CC claimed. The DNA can be used to make the hormone receptor proteins and  
 CC functional modified forms in quantities not previously possible. The  
 CC receptor proteins can be used to screen cpds. for receptor-agonist or  
 CC receptor-antagonist activity. They can also be used in diagnostic assays.  
 SQ Sequence 456 AA;

Query Match 92.3%; Score 36; DB 1; Length 456;  
 Best Local Similarity 83.3%; Pred. NO. 37;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 159 CRFKKC 164

## RESULT 11

ID R26899 standard; protein; 490 AA.  
 AC R26899; 1993 (first entry)  
 DT 17-FEB-1993 (first entry)  
 DE Herba-T sequence.  
 KW Thyroid hormone; receptor; antibodies; immunoassay; oncogene;  
 KW tumour; cancer; therapy.  
 OS Homo sapiens.  
 PN US5144007-A.  
 PD 01-SEP-1992.  
 PF 03-NOV-1988; 266529.  
 PR 03-NOV-1988; US-266529.  
 PR 24-OCT-1990; US-604334.  
 PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 PI Pfehl M;  
 DR WPI: 92-315568/38.

DR N-PSDB: Q28370.  
 PT Thyroid hormone receptor herba-T and its antibodies - bind  
 PT thyroid hormone, especially I3, for assessing the presence of  
 PT receptor in tissues e.g. in thyroid diseases  
 PS Claim 2; Fig 1; 16pp; English.  
 CC The human thyroid hormone receptor (herba-T) protein sequence was  
 CC deduced from the cDNA sequence obtd. by screening a human testis  
 CC cDNA library with a labelled probe obtd. from a randomly primed  
 CC and transcribed 600 bp AVAL-PSTI fragment from v-herba encoding  
 CC the cysteine rich domain. The herba-T polypeptide may be used to  
 CC generate monoclonal and polyclonal antibodies to herba-T which may  
 CC be employed in immunoassays. Since herba-T is a human homologue of  
 CC the viral oncogene herba, tumour growth can be affected by the level  
 CC of herba-T. Measurement of herba-T or its RNA will help assess this.  
 CC The protein and cDNA are also useful for the identification of  
 CC optimal agonists and antagonists of herba-T for therapeutic use, e.g.  
 CC to develop tissue specific thyroid hormone therapies for certain  
 CC diseases and cancers.  
 SQ Sequence 490 AA;

Query Match 92.3%; Score 36; DB 1; Length 490;  
 Best Local Similarity 83.3%; Pred. NO. 39;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 110 CRFKKC 115

## RESULT 12

ID R47621 standard; protein; 514 AA.  
 AC R47621;  
 DT 07-JUL-1994 (first entry)  
 DE Rat thyroid hormone receptor.  
 KW Pituitary-specific thyroid hormone receptor; PSTHR; erba gene;  
 KW hypothyroidism; hyperthyroidism.  
 OS Rattus sp.  
 PN US5284999-A.  
 PD 08-FEB-1994.  
 PF 05-APR-1990; 504806.  
 PR 05-APR-1990; US-504806.  
 PR 13-MAY-1993; US-069643.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PI Chin WW, Hodin RA, Lazar MA;  
 DR WPI: 94-056439/07.  
 DR P-PSDB: Q54086.  
 PT Recombinant DNA molecule encodes pituitary-specific thyroid  
 PT hormone receptor - useful for therapy of e.g. hypo- and  
 PT hyper-thyroidism  
 PS Disclosure; Column 21-22; 12pp; English.  
 CC A rat GH3 cell cDNA library was screened with rat erba sequences to  
 CC isolate a gene encoding PSTHR. The sequence of the full-length cDNA  
 CC for PSTHR is given in Q54086, and the deduced amino acid sequence in  
 CC R47621.  
 SQ Sequence 514 AA;

Query Match 92.3%; Score 36; DB 1; Length 514;  
 Best Local Similarity 83.3%; Pred. NO. 40;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 I:|||||  
 Db 217 CRFKKC 222

## RESULT 13

ID R71565 standard; protein; 579 AA.  
 AC R71565;  
 DT 31-OCT-1995 (first entry)



DE EAR-1r gene product (nuclear receptor) from human osteosarcoma cells.  
 KW nuclear receptor protein; osteosarcoma; EAR-1r; generation; growth;  
 KW differentiation; human; erb-A related gene.  
 OS Homo sapiens.  
 PN J0703798-A.  
 PD 03-FEB-1995.  
 PF 20-JUL-1993; 200333.  
 PR 20-JUL-1993; JP-200333.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 DR WPI; 95-110650/15.  
 DR N-PSDB; Q85975.  
 PT Nuclear receptor protein derived from osteo:blastic osteo:sarcoma  
 PT - encoded by the EAR-1r gene, for investigation of generation of  
 PT fat soluble active substances  
 PS Claim 1; Page 4-6; 8pp; Japanese.  
 CC The EAR-1r gene (human erb-A related gene) (Q85975) encodes a nuclear  
 CC receptor protein (R71565) derived from human osteosarcoma (HOS)  
 CC cells. The expression of EAR-1r is most abundant in the brain and  
 CC was detectedd slightly in heart and placenta. The receptor protein  
 CC may be used for the investigation of generation, differentiation  
 CC and growth of fat soluble physiological substances.  
 SQ Sequence 579 AA;

Query Match 92.3%; Score 36; DB 1; Length 579;  
 Best Local Similarity 83.3%; Pred. No. 44;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
 Db 159 CRFKKC 164  
 I:||||

## RESULT 14

R78318  
 ID R78318 standard; Protein; 410 AA.

AC R78318;  
 DE 06-MAR-1996 (first entry)  
 DT Human thyroid hormone receptor alpha-1.  
 KW Human; thyroid hormone; receptor; alpha-1; htr-alpha-1; analysis;  
 KW assays; therapeutic agents; antibodies; affinity purification;  
 KW detection; quantification.  
 OS Homo sapiens.  
 PN US5438126-A.  
 PD 01-AUG-1995.  
 PF 11-SEP-1989; 405342.  
 PR 11-SEP-1989; US-405342.  
 PR 03-FEB-1992; US-830766.  
 PA (ARCH-) ARCH DEV CORP.  
 PI Degroot LJ, Nakai A;  
 DR WPI; 95-274923/36.  
 DR N-PSDB; Q95110.  
 PT New human thyroid hormone receptor nucleic acid - used to develop  
 PT prods. for use in analysis, assays and therapeutic agents  
 PS Claim 1; Fig 1; 7pp; English.  
 CC Q95110 encodes R78318 human thyroid hormone receptor alpha-1 (htr-  
 CC alpha-1). The nucleic acid can be used to develop prods. for use in  
 CC analysis assays and therapeutic agents. htr-alpha-1 can be used in  
 CC tests for thyroid function, or for producing antibodies for use in  
 CC affinity purificn., detection and quantification.  
 SQ Sequence 410 AA;

Query Match 92.3%; Score 36; DB 1; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
 Db 110 CRFKKC 115  
 I:||||

## RESULT 15

W41493  
 ID W41493 standard; Protein; 410 AA.  
 AC W41493;  
 DE 22-JUN-1998 (first entry)  
 DT LambdaMe2 containing human thyroid hormone receptor protein htr alpha 1.  
 KW Human; thyroid hormone receptor protein; htr alpha 1; lambdaMe2.  
 OS Homo sapiens.  
 FH Key  
 FT Misc\_difference 139 Location/Qualifiers  
 FT /note= "encoded by GAG"  
 FT Misc\_difference 245 /note= "encoded by GAA"  
 FT Misc\_difference 285 /note= "encoded by GAG"

PN US5712372-A.  
 PD 27-JAN-1998.  
 PF 21-MAR-1995; 407909.  
 PR 11-SEP-1989; US-405342.  
 PR 03-FEB-1992; US-830766.  
 PR 21-MAR-1995; US-407909.  
 PA (ARCH-) ARCH DEV CORP.  
 PI Degroot LJ, Nakai A;  
 DR WPI; 98-120030/11.  
 DR N-PSDB; V17551.

PT Thyroid hormone receptor polypeptide - and antibody to it, useful in  
 PT testing thyroid function  
 PS Claim 1; Fig 1; 8pp; English.  
 CC The present sequence represents a recombinant form of human thyroid  
 CC hormone receptor protein htr alpha 1. The present invention also  
 CC describes an antibody that is immunoreactive with the protein. The  
 CC protein is expected to be useful for therapeutic purposes, for producing  
 CC antibodies and for developing more effective tests of thyroid function.  
 CC The antibodies can be used for affinity purification of htr alpha 1.  
 SQ Sequence 410 AA;

Query Match 92.3%; Score 36; DB 1; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
 Db 110 CRFKKC 115  
 I:||||

Search completed: September 7, 1999, 20:37:16  
 Job time: 18465 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:36 ; Search time 80.79 Seconds  
(without alignments)  
0.733 Million cell updates/sec

Title: US-09-124-280A-14  
Perfect score: 39  
Sequence: 1 CKFKKC 6  
Scoring table: BLOSUM62  
Searched: 106577 seqs, 9868381 residues  
Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	6	1	US-08-097-830E-14
2	39	100.0	6	2	US-08-456-112B-14
3	32	82.1	439	1	US-08-333-358-10
4	32	82.1	6	1	US-08-097-830E-32
5	32	82.1	439	1	US-08-463-694-10
6	32	82.1	439	1	US-08-694-501-10
7	32	82.1	225	2	US-08-462-169B-22
8	32	82.1	475	2	US-08-484-200-2
9	32	82.1	440	2	US-08-484-200-4
10	30	76.9	795	1	US-07-716-827C-5
11	30	76.9	82	1	US-07-965-674-1
12	30	76.9	83	1	US-07-965-674-2
13	30	76.9	84	1	US-07-965-674-3
14	30	76.9	83	1	US-07-965-674-4
15	30	76.9	89	1	US-08-087-151-3
16	30	76.9	1400	1	US-08-080-255-7
17	30	76.9	89	1	US-08-561-072-3
18	30	76.9	10	1	US-08-097-830E-18
19	30	76.9	468	1	US-08-459-287-2
20	30	76.9	468	1	US-08-459-287-3
21	30	76.9	414	1	US-08-255-471-9
22	30	76.9	10	2	US-08-456-112B-18
23	30	76.9	43	3	PCT-US93-05640-18
24	30	76.9	1400	3	PCT-US93-05857-7
25	30	76.9	82	3	PCT-US93-09523-1
26	30	76.9	83	3	PCT-US93-09523-2
27	30	76.9	84	3	PCT-US93-09523-3
28	30	76.9	83	3	PCT-US93-09523-4
29	29	74.4	533	1	US-07-952-800-2
30	29	74.4	446	1	US-07-952-800-4
31	29	74.4	513	1	US-08-464-266-2
32	29	74.4	513	1	US-08-464-272-2
33	29	74.4	462	1	US-08-336-408B-2
34	29	74.4	467	1	US-08-336-408B-4
35	29	74.4	463	1	US-08-336-408B-6
36	29	74.4	412	2	US-08-313-288B-18
37	29	74.4	15	2	US-08-783-275-8
38	29	74.4	15	2	US-08-727-708-8
39	29	74.4	389	2	US-08-485-449-2

40 29 74.4 376 2 US-08-485-449-5  
41 29 74.4 389 2 US-08-485-449-6  
42 29 74.4 389 2 US-08-485-449-7  
43 29 74.4 462 3 PCT-US91-00399-2  
44 29 74.4 467 3 PCT-US91-00399-4  
45 29 74.4 463 3 PCT-US91-00399-6

ALIGNMENTS

RESULT 1  
US-08-097-830E-14  
; Sequence 14, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and  
; OTHER INFORMATION: Cys  
; US-08-097-830E-14

Query Match 100.0%; Score 39; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CKFKKC 6  
DB 1 CKFKKC 6  
RESULT 2  
US-08-456-112B-14  
; Sequence 14, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-14

Query Match 100.0%; Score 39; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04; Indels 0;  
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 CKFKKC 6  
|:|:|:|  
Db 1 CKFKKC 6

RESULT 3  
US-08-333-358-10  
Sequence 10, Application US/08333358  
Patent No. 5571696  
GENERAL INFORMATION:  
APPLICANT: EVANS Ph.D., RONALD M.  
APPLICANT: MANGELSDORF Ph.D., DAVID J.  
APPLICANT: ONG Ms., ESTELITA S.  
APPLICANT: ORO Ph.D., ANTHONY E.  
APPLICANT: BORGMEYER Ph.D., UWE K.  
APPLICANT: GIGUERE Ph.D., VINCENT NMN  
APPLICANT: YAO Mr., TSO-PANG NMN  
TITLE OF INVENTION: NOVEL RECEPTORS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 So. Flower St., Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: US  
ZIP: 90071-2921  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,358  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/761,068  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter Ph.D., Stephen E.  
REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P31 8936  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 439 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-358-10

Query Match 82.1%; Score 32; DB 1; Length 439;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFKKC 6  
|:|:|:|  
Db 127 CRFKQC 132

RESULT 4  
US-08-097-830E-32  
Sequence 32, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:  
OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-32

Query Match 82.1%; Score 32; DB 1; Length 6;  
 Best Local Similarity 83.3%; Pred. No. 7.9e+04;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 ||| ||  
 Db 1 CKELKC 6

## RESULT 5

US-08-463-694-10  
 ; Sequence 10, Application US/08463694  
 ; Patent No. 5696233  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EVANS Ph.D., RONALD M.  
 ; APPLICANT: MANGELSDORF Ph.D., DAVID J.  
 ; APPLICANT: ONG Ms., ESTELITA S.  
 ; APPLICANT: ORO Ph.D., ANTHONY E.  
 ; APPLICANT: BORGMEYER Ph.D., UWE K.  
 ; APPLICANT: GIGUERE Ph.D., VINCENT NMN  
 ; APPLICANT: YAO Mr., TSO-PANG NMN  
 ; TITLE OF INVENTION: NOVEL RECEPTORS  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
 ; STREET: 444 So. Flower St., Suite 2000  
 ; CITY: Los Angeles  
 ; STATE: CA  
 ; COUNTRY: US  
 ; ZIP: 90071-2921

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/463,694  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/761,068  
 FILING DATE: 17-SEP-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Reiter Ph.D., Stephen E.  
 REGISTRATION NUMBER: 31192  
 REFERENCE/DOCKET NUMBER: P31 8936  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 439 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-463-694-10

Query Match 82.1%; Score 32; DB 1; Length 439;  
 Best Local Similarity 86.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 ||| ||  
 Db 127 CRQKC 132

## RESULT 6

US-08-694-501-10  
 ; Sequence 10, Application US/08694501  
 ; Patent No. 5710004  
 ; GENERAL INFORMATION:

APPLICANT: EVANS Ph.D., RONALD M.  
 APPLICANT: MANGELSDORF Ph.D., DAVID J.  
 APPLICANT: ONG Ms., ESTELITA S.  
 APPLICANT: ORO Ph.D., ANTHONY E.  
 APPLICANT: BORGMEYER Ph.D., UWE K.  
 APPLICANT: GIGUERE Ph.D., VINCENT NMN  
 APPLICANT: YAO Mr., TSO-PANG NMN  
 TITLE OF INVENTION: NOVEL RECEPTORS  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
 STREET: 444 So. Flower St., Suite 2000  
 CITY: Los Angeles  
 STATE: CA  
 COUNTRY: US  
 ZIP: 90071-2921

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/694,501  
 FILING DATE: 07-AUG-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/333,358  
 FILING DATE:  
 APPLICATION NUMBER: US/07/761,068  
 FILING DATE:

ATTORNEY/AGENT INFORMATION:  
 NAME: Reiter Ph.D., Stephen E.  
 REGISTRATION NUMBER: 31192  
 REFERENCE/DOCKET NUMBER: P31 8936  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 439 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-694-501-10

Query Match 82.1%; Score 32; DB 1; Length 439;  
 Best Local Similarity 86.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 ||| ||  
 Db 127 CRQKC 132

## RESULT 7

US-08-462-169B-22  
 ; Sequence 22, Application US/08462169B  
 ; Patent No. 5773252  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John Greene and Craig A. Rosen  
 ; TITLE OF INVENTION: Fibroblast Growth Factor-15  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
 ; STREET: 6 BECKER FARM ROAD  
 ; CITY: ROSELAND  
 ; STATE: NEW JERSEY  
 ; COUNTRY: USA  
 ; ZIP: 07068  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462.169B  
FILING DATE: 05 JUN 95  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-441 (PF203)  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 225 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-462-169B-22

Query Match 82.1%; Score 32; DB 2; Length 225;  
Best Local Similarity 66.7%; Pred. No. 72;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
Db 143 CRFKC 148  
|:|:|

RESULT 8  
US-08-484-200-2  
; Sequence 2, Application US/08484200  
; Patent No. 5861274  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: FORMAN, BARRY M.  
; APPLICANT: KLIWER, STEVEN A.  
; APPLICANT: ONG, ESTELITA S.  
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9971  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-1995  
; TELEFAX: 619-546-9392  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-484-200-2

Query Match 82.1%; Score 32; DB 2; Length 475;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
Db 163 CRFKC 168  
|:|:|

RESULT 9  
US-08-484-200-4  
; Sequence 4, Application US/08484200  
; Patent No. 5861274  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: FORMAN, BARRY M.  
; APPLICANT: KLIWER, STEVEN A.  
; APPLICANT: ONG, ESTELITA S.  
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9971  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-1995  
; TELEFAX: 619-546-9392  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 440 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-484-200-4

Query Match 82.1%; Score 32; DB 2; Length 440;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
Db 127 CRFKC 132  
|:|:|

RESULT 10  
US-07-716-827C-5  
; Sequence 5, Application US/07716827C  
; Patent No. 5215916  
; GENERAL INFORMATION:  
; APPLICANT: Simons Jr., Stoney S.  
; APPLICANT: Yamamoto, K. R.

APPLICANT: Chakraborti, P. K.  
APPLICANT: Garabedian, M. J.  
TITLE OF INVENTION: SUPER GLUCOCORTICOID RECEPTORS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman, Darby & Cushman  
STREET: Eleventh floor, 1615 L Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036-5601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/716,827C  
FILING DATE: 19910619  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/84453  
TELEPHONE: (202)8613000  
TELEFAX: (202)822-8944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 795 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-716-827C-5

Query Match 76.9%; Score 30; DB 1; Length 795;  
Best Local Similarity 50.0%; Pred. No. 3.7e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
Db 495 CRYKNC 500

RESULT 11  
US-07-965-674-1  
Sequence 1, Application US/07965674  
Patent No. 5380646  
GENERAL INFORMATION:  
APPLICANT: Knight, Linda C.  
TITLE OF INVENTION: Thrombus Detection Using  
STREET: 406 University Services Building  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/965,674  
FILING DATE: 19921019  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-173  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5380646e  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 82 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-965-674-1

Query Match 76.9%; Score 30; DB 1; Length 82;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKK 5  
Db 50 CKFKK 54

RESULT 12  
US-07-965-674-2  
Sequence 2, Application US/07965674  
Patent No. 5380646  
GENERAL INFORMATION:  
APPLICANT: Knight, Linda C.  
TITLE OF INVENTION: Thrombus Detection Using  
STREET: 406 University Services Building  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/965,674  
FILING DATE: 19921019  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-173  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5380646e  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
US-07-965-674-2

Query Match 76.9%; Score 30; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKK 5  
 Db 51 CKFKK 55

## RESULT 13

US-07-965-674-3  
 ; Sequence 3, Application US/07965674  
 ; Patent No. 5380646  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knight, Linda C.  
 ; TITLE OF INVENTION: Thrombus Detection Using  
 ; TITLE OF INVENTION: Radiolabelled Disintegrins  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Temple University - Of The Common-  
 ; ADDRESSEE: wealth System of Higher Education  
 ; STREET: 406 University Services Building  
 ; CITY: Philadelphia  
 ; STATE: Pennsylvania  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19122

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WordPerfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/965,674  
 ; FILING DATE: 19921019  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Monaco, Daniel A.  
 ; REGISTRATION NUMBER: 30,480  
 ; REFERENCE/DOCKET NUMBER: 6056-173  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (215) 568-8383  
 ; TELEFAX: (215) 568-5549  
 ; TELEX: No. 5380646e  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 84 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: linear  
 ; US-07-965-674-3

Query Match 76.9%; Score 30; DB 1; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKK 5  
 Db 52 CKFKK 56

## RESULT 14

US-07-965-674-4  
 ; Sequence 4, Application US/07965674  
 ; Patent No. 5380646  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Knight, Linda C.  
 ; TITLE OF INVENTION: Thrombus Detection Using  
 ; TITLE OF INVENTION: Radiolabelled Disintegrins  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Temple University - Of The Common-  
 ; ADDRESSEE: wealth System of Higher Education

STREET: 406 University Services Building  
 CITY: Philadelphia  
 STATE: Pennsylvania  
 COUNTRY: U.S.A.  
 ZIP: 19122  
 COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WordPerfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/965,674  
 ; FILING DATE: 19921019  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Monaco, Daniel A.  
 ; REGISTRATION NUMBER: 30,480  
 ; REFERENCE/DOCKET NUMBER: 6056-173  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (215) 568-8383  
 ; TELEFAX: (215) 568-5549  
 ; TELEX: No. 5380646e  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 83 amino acids  
 ; TYPE: AMINO ACID  
 ; TOPOLOGY: linear  
 ; US-07-965-674-4

Query Match 76.9%; Score 30; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKK 5  
 Db 51 CKFKK 55

## RESULT 15

US-08-087-151-3  
 ; Sequence 3, Application US/08087151  
 ; Patent No. 5468624  
 ; GENERAL INFORMATION:  
 ; APPLICANT: E. Brad Thompson  
 ; APPLICANT: Lynne V. Nazareth  
 ; TITLE OF INVENTION: ANTI-TUMOR ACTIVITY OF A MODIFIED FRAGMENT  
 ; TITLE OF INVENTION: OF THE GLUCOCORTICOID RECEPTOR  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold, White & Durkee  
 ; STREET: P.O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77210  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WordPerfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/087,151  
 ; FILING DATE: July 7, 1993  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hodgins, Daniel S.  
 ; REGISTRATION NUMBER: 31,026  
 ; REFERENCE/DOCKET NUMBER: UTSG:141/HOD  
 ; TELECOMMUNICATION INFORMATION:



```

; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: No. 5468624 Applicable
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-087-151-3

```

```

Query Match      76.9%; Score 30; DB 1; Length 89;
Best Local Similarity 50.0%; Pred. NO. 73;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CKFKC 6
Db 79 CRYKC 84

```

```

Search completed: September 7, 1999, 22:38:37
Job time: 7921 sec

```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:28 ; Search time 116.8 Seconds  
(without alignments)  
2.058 Million cell updates/sec

Title: US-09-124-280A-14

Perfect score: 39

Sequence: 1 CKFKKC 6

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	92.3	456	1 TVCHUR	thyroid hormone re
2	36	92.3	369	1 TVCHTB	thyroid hormone re
3	36	92.3	454	1 TVRTAR	thyroid hormone re
4	36	92.3	410	1 QRMSAI	thyroid hormone re
5	36	92.3	418	1 TVXITA	thyroid hormone re
6	36	92.3	408	1 TVCHVR	thyroid hormone re
7	36	92.3	398	1 TVFVVR	transforming prote
8	36	92.3	560	2 S27874	steroid hormone re
9	36	92.3	406	2 A44776	transforming prote
10	36	92.3	418	2 B36067	thyroid hormone re
11	36	92.3	492	2 S06907	thyroid hormone re
12	36	92.3	461	2 A40377	thyroid hormone re
13	36	92.3	373	2 D36067	thyroid hormone re
14	36	92.3	373	2 C36067	thyroid hormone re
15	36	92.3	387	2 S35744	transforming prote
16	36	92.3	410	2 A40917	thyroid hormone re
17	36	92.3	461	2 A31820	thyroid hormone re
18	36	92.3	410	2 S06410	thyroid hormone re
19	36	92.3	410	2 S09178	thyroid hormone re
20	36	92.3	475	2 A57035	thyroid hormone be
21	36	92.3	490	2 A30893	thyroid hormone re
22	36	92.3	373	2 I51165	gene c-erbA-beta p
23	36	92.3	492	2 I57696	c-erbA-alpha-2-rel
24	36	92.3	416	2 I51096	thyroid hormone re
25	36	92.3	391	2 I51097	thyroid hormone re
26	36	92.3	452	2 S14417	thyroid hormone re
27	36	92.3	410	2 S14416	thyroid hormone re
28	36	92.3	369	2 S58211	beta-thyroid hormo
29	36	92.3	413	2 S14418	thyroid hormone re
30	36	92.3	68	2 S06163	transcription fact
31	36	92.3	342	2 A37952	probable thyroid h
32	36	92.3	113	2 S43435	nuclear receptor p
33	36	92.3	864	2 A49070	ecdysone-inducible
34	36	92.3	556	2 I45066	steroid hormone re
35	36	92.3	579	2 A57057	orphan nuclear hor
36	36	92.3	614	2 A32608	thyroid hormone re
37	36	92.3	508	2 A30226	thyroid/steroid ho
38	36	92.3	576	2 A57048	nuclear receptor R
39	36	92.3	578	2 S52913	nuclear receptor R

ALIGNMENTS

RESULT 1

TVHUR

thyroid hormone receptor beta - human

N:Alternate names: erba transforming protein

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 05-Sep-1997

C:Accession: A25237; I52900

R:Weinberger, C.; Thompson, C.C.; Ong, E.S.; Lebo, R.; Gruel, D.J.; Evans, R.M.

Nature 324, 641-646, 1986

A:Title: The c-erb-A gene encodes a thyroid hormone receptor.

A:Reference number: A25237; MUID:87090375

A:Accession: A25237

A:Molecule type: mRNA

A:Residues: 1-456 <WE1>

A:Cross-references: EMBL:X04707; NID:g31206; PID:g31207

R:Weinberger, C.; Giguere, V.; Hollenberg, S.; Rosenfeld, M.G.; Evans, R.M.

Cold Spring Harb. Symp. Quant. Biol. 51, 759-772, 1988

A:Title: Human steroid receptors and erba proto-oncogene products: Members of a new s

A:Reference number: I52900; MUID:87217099

A:Accession: I52900

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-456 <WE2>

A:Cross-references: GB:M26747; NID:g180252; PID:g180253

C:Genetics: GDB:THRB; ERBA2

A:Gene: GDB:THRB; ERBA2

A:Map position: 3p24.1-3p22

C:Superfamily: thyroid hormone receptor; erba transforming protein homology

C:Keywords: DNA binding; nucleus; proto-oncogene; thyroid hormone receptor; transcript

F:100-376/Domain: erba transforming protein homology <ERBA>

F:102-122/Region: zinc finger CCCC motif

F:140-164/Region: zinc finger CCCC motif

F:239-456/Domain: hormone binding #status predicted <HR>

Query Match 92.3%; Score 36; DB 1; Length 456;

Best Local Similarity 83.3%; Pred. No. 12;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6

|:||||

Db 159 CKFKKC 164

RESULT 2

TVCHTB

thyroid hormone receptor beta - chicken

N:Alternate names: erba transforming protein

C:Species: Gallus gallus (chicken)

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Sep-1997

C:Accession: S09625; A38460

R:Forrest, D.; Stoeberg, M.; Vennstroem, B.

EMBO J. 9, 1519-1528, 1990

A:Title: Contrasting developmental and tissue-specific expression of alpha and beta t

A:Reference number: S09625; MUID:90228351

A:Accession: S09625

A:Molecule type: mRNA

A:Residues: 1-369 <FOR>

A:Cross-references: EMBL:X17504

A:Experimental source: kidney

R:Showers, M.O.; Darling, D.S.; Kieffer, G.D.; Chin, W.W.

DNA Cell Biol. 10, 211-221, 1991  
A:Title: Isolation and characterization of a cDNA encoding a chicken beta thyroid hormone receptor  
A:Reference number: A38460; MUID:91190272  
A:Accession: A38460  
A:Molecule type: mRNA  
A:Residues: 1-116, 'V', 118-143, 'V', 145-201, 'L', 203-214, 'V', 216-369 <SHO>  
A:CROSS-references: GB:M65207; NID:g212801; PID:g212802  
A:Experimental source: liver  
A:Note: the authors translated the codon GGT for residue 117 as Gly, GTA for residue 144 as Val, and GGT for residue 214 as Val.  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; proto-oncogene; thyroid hormone receptor; transcriptid  
F:13-289/Domain: erba transforming protein homology <ERBA>  
F:14-82/Domain: DNA binding #status predicted <DNA>  
F:15-37/Region: zinc finger CCCC motif  
F:53-77/Region: zinc finger CCCC motif  
F:132-369/Domain: hormone binding #status predicted <HRB>

Query Match 92.3%; Score 36; DB 1; Length 369;  
Best Local Similarity 83.3%; Pred. No. 10;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 72 CRFKKC 77

RESULT 3  
TVRTAR  
N:Thyroid hormone receptor alpha-2 - rat  
N:Alternate names: transforming protein (erba) homolog  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 05-Sep-1997  
C:Accession: A36752; A40923; S00927; A35581  
R:Nikodem, V.  
submitted to the EMBL Data Library, April 1988  
A:Reference number: A36752  
A:Accession: A36752  
A:Molecule type: mRNA  
A:Residues: 1-454 <NIK>  
A:CROSS-references: EMBL:X07409; NID:g57390; PID:g57391  
R:Lazar, M.A.; Chin, W.W.  
Mol. Endocrinol. 2, 479-484, 1988  
A:Title: Regulation of two c-erbA messenger ribonucleic acids in rat GH-3 cells by thyroid hormone  
A:Reference number: A40923; MUID:88334548  
A:Accession: A40923  
A:Molecule type: mRNA  
A:Residues: 82-443, 'C', 445-454 <LAZ>  
A:CROSS-references: GB:M31177; NID:g204049; PID:g204050  
R:Mitsushashi, T.; Tenmyson, G.; Nikodem, V.  
Nucleic Acids Res. 16, 5697, 1988  
A:Title: Nucleotide sequence of novel cDNAs generated by alternative splicing of a rat thyroid hormone receptor  
A:Reference number: S00927; MUID:88262581  
A:Accession: S00927  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 333-454 <MIT>  
R:Mitsushashi, T.; Tenmyson, G.E.; Nikodem, V.M.  
Proc. Natl. Acad. Sci. U.S.A. 85, 5804-5808, 1988  
A:Title: Alternative splicing generates messages encoding rat c-erbA proteins that do not bind thyroid hormone  
A:Reference number: A35581; MUID:88320332  
A:Accession: A35581  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 333-454 <MIT>  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; nucleus; proto-oncogene; thyroid hormone  
F:13-289/Domain: erba transforming protein homology <ERBA>  
F:15-35/Region: zinc finger CCCC motif  
F:53-77/Region: zinc finger CCCC motif

Query Match 92.3%; Score 36; DB 1; Length 454;

Best Local Similarity 83.3%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 72 CRFKKC 77

RESULT 4  
QRMSA1  
thyroid hormone receptor alpha-1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 05-Sep-1997  
C:Accession: S14690; S08690  
R:Masuda, M.; Yasuhara, S.; Yamashita, M.; Shibuya, M.; Odaka, T.  
Nucleic Acids Res. 18, 3055, 1990  
A:Title: Nucleotide sequence of the murine thyroid hormone receptor (alpha-1) cDNA.  
A:Reference number: S14690; MUID:90272421  
A:Accession: S14690  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-410 <MAS>  
A:CROSS-references: EMBL:X51983; NID:g50385; PID:g50386  
C:Genetics:  
A:Gene: erba-alpha  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; proto-oncogene; thyroid hormone receptor  
F:51-327/Domain: erba transforming protein homology <ERBA>  
F:53-73/Region: zinc finger CCCC motif  
F:91-115/Region: zinc finger CCCC motif  
F:190-410/Domain: hormone binding #status predicted <HRB>

Query Match 92.3%; Score 36; DB 1; Length 410;  
Best Local Similarity 83.3%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
|:|||||  
Db 110 CRFKKC 115

RESULT 5  
TVXLTA  
thyroid hormone receptor alpha-1 - African clawed frog  
N:Alternate names: thyroid hormone receptor alpha-A  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Sep-1997  
C:Accession: S09605; A36067  
R:Brooks, A.R.; Sweeney, G.; Old, R.W.  
Nucleic Acids Res. 17, 9395-9405, 1989  
A:Title: Structure and functional expression of a cloned Xenopus thyroid hormone receptor  
A:Reference number: S09605; MUID:90067936  
A:Accession: S09605  
A:Molecule type: mRNA  
A:Residues: 1-418 <BRO>  
A:CROSS-references: EMBL:X17385; NID:g65141; PID:g65142  
R:Yaoita, Y.; Shi, Y.; Brown, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7090-7094, 1990  
A:Title: Xenopus laevis alpha and beta thyroid hormone receptors.  
A:Reference number: A36067; MUID:90384953  
A:Accession: A36067  
A:Molecule type: mRNA  
A:Residues: 1-53, 'V', 55-136, 'G', 138-418 <YAO>  
A:CROSS-references: GB:M35343; NID:g214829; PID:g214830  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; thyroid hormone receptor; transcription regulation;  
F:59-335/Domain: erba transforming protein homology <ERBA>  
F:60-128/Domain: DNA binding #status predicted <DNA>  
F:61-81/Region: zinc finger CCCC motif  
F:99-123/Region: zinc finger CCCC motif  
F:178-415/Domain: hormone binding #status predicted <HRB>

Query Match 92.3%; Score 36; DB 1; Length 418;  
 Best Local Similarity 83.3%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|||||  
 Db 118 CKFKKC 123

## RESULT 6

TVCHVR

thyroid hormone receptor alpha - chicken

N;Alternate names: erba transforming protein

C:Species: Gallus gallus (chicken)

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 05-Sep-1997

C:Accession: A25236; S00029

R:Sap, J.; Munoz, A.; Damm, K.; Goldberg, Y.; Ghysdael, J.; Leutz, A.; Beug, H.; Vennstr

Nature 324, 635-640, 1986

A;Title: The c-erb-A protein is a high-affinity receptor for thyroid hormone.

A:Reference number: A25236; MUID:87090374

A:Accession: A25236

A:Molecule type: mRNA

A:Residues: 1-408 &lt;SAP&gt;

A:Cross-references: GB:Y00987; NID:963177; PID:963178

R:Zahraoui, A.; Cuny, G.

Eur. J. Biochem. 166, 63-69, 1987

A:Title: Nucleotide sequence of the chicken proto-oncogene c-erba corresponding to domain

A:Reference number: S00029; MUID:87246693

A:Accession: S00029

A:Molecule type: DNA

A:Residues: 1-239 &lt;ZAH&gt;

A:Cross-references: GB:X04858

C:Genetics:

A:Gene: erba

C:Superfamily: thyroid hormone receptor; erba transforming protein homology

C:Keywords: DNA binding; nucleus; phosphoprotein; proto-oncogene; thyroid hormone recept

F:49-325/Domain: erba transforming protein homology &lt;ERBA&gt;

F:51-118/Domain: DNA binding #status predicted &lt;DNA&gt;

F:51-71/Region: zinc finger CCCC motif

F:89-113/Region: zinc finger CCCC motif

F:168-408/Domain: hormone binding #status predicted &lt;HRB&gt;

Query Match 92.3%; Score 36; DB 1; Length 408;  
 Best Local Similarity 83.3%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|||||  
 Db 108 CKFKKC 113

## RESULT 7

TVFVR

transforming protein erba - avian erythroblastosis virus

C:Species: avian erythroblastosis virus

A:Note: host Gallus gallus (chicken)

C:Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 16-Feb-1997

C:Accession: A03248

R:Debulre, B.; Henry, C.; Benaisa, M.; Biserte, G.; Claverie, J.M.; Saule, S.; Martin,

Science 224, 1456-1459, 1984

A;Title: Sequencing the erba gene of avian erythroblastosis virus reveals a new type of

A:Reference number: A38022; MUID:84223957

A:Accession: A03248

A:Molecule type: genomic RNA

A:Residues: 1-398 &lt;DEB&gt;

A:Cross-references: GB:K02006

C:Comment: This protein is synthesized as a gag-erba polyprotein.

C:Genetics:

A:Gene: erba

C:Superfamily: thyroid hormone receptor; erba transforming protein homology

C:Keywords: DNA binding; oncogene; polyprotein; transcription regulation; transforming p

F:35-311/Domain: erba transforming protein homology <ERBA>  
 F:37-57/Region: zinc finger CCCC motif  
 F:75-99/Region: zinc finger CCCC motif

Query Match 92.3%; Score 36; DB 1; Length 398;  
 Best Local Similarity 83.3%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|||||  
 Db 94 CKFKKC 99

## RESULT 8

S27874

steroid hormone receptor homolog LRH-1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 10-Oct-1997

C:Accession: S27874

R:Tugwood, J.D.; Issemann, I.; Green, S.

submitted to the EMBL Data Library, February 1992

A:Description: LRH-1: A nuclear hormone receptor active in the absence of exogenous 1

A:Reference number: S27873

A:Accession: S27874

A:Molecule type: mRNA

A:Residues: 1-560 &lt;TUG&gt;

A:Cross-references: EMBL:M81385; NID:g198872; PID:g198875

C:Function:

A:Description: transcription regulation; probably a steroid hormone receptor

C:Superfamily: steroid hormone receptor Ad4BP; erba transforming protein homology

C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription factor; zin

F:105-478/Domain: erba transforming protein homology &lt;ERBA&gt;

F:107-127/Region: zinc finger CCCC motif

F:143-167/Region: zinc finger CCCC motif

Query Match 92.3%; Score 36; DB 2; Length 560;  
 Best Local Similarity 83.3%; Pred. No. 14;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|||||  
 Db 162 CKFKKC 167

## RESULT 9

A44776

transforming protein (erba) - avian erythroblastosis virus (strain ES4)

C:Species: avian erythroblastosis virus

C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 03-Jun-1996

C:Accession: A44776

R:Bruskin, A.; Jackson, J.; Bishop, J.M.; McCarley, D.J.; Schatzman, R.C.

Oncogene 5, 15-24, 1990

A;Title: Six amino acids from the retroviral gene gag greatly enhance the transformin

A:Reference number: A44776; MUID:90206603

A:Accession: A44776

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-406 &lt;BRU&gt;

A:Cross-references: GB:X52211

C:Genetics:

A:Gene: erba

C:Superfamily: thyroid hormone receptor; erba transforming protein homology

C:Keywords: zinc finger

F:36-332/Domain: erba transforming protein homology &lt;ERBA&gt;

F:58-78/Region: zinc finger

F:96-120/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 406;  
 Best Local Similarity 83.3%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
|:|||||  
Db 115 CRFKKC 120

RESULT 10  
B36067  
thyroid hormone receptor alpha-B - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 08-Sep-1997  
C:Accession: B36067  
R:Yaota, Y.; Shi, Y.; Brown, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7090-7094, 1990  
A:Title: Xenopus laevis alpha and beta thyroid hormone receptors.  
A:Reference number: A36067; MUID:90384953  
A:Accession: B36067  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-418 <YAO>  
A:Cross-references: GB:M35344; NID:g214831; PID:g214832  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; thyroid hormone receptor; transcription regulation; zinc finger  
F:59-335/Domain: erba transforming protein homology <ERBA>  
F:61-81/Region: zinc finger  
F:99-123/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 418;  
Best Local Similarity 83.3%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
|:|||||  
Db 118 CRFKKC 123

RESULT 11  
S06907  
thyroid hormone receptor alpha-2 (clone rTRalpa2) - rat  
N:Alternate names: transforming protein (erba) homolog  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 16-Feb-1997  
C:Accession: S06907  
R:Izumo, S.; Mahdavi, V.  
Nature 334, 539-542, 1988  
A:Title: Thyroid hormone receptor alpha isoforms generated by alternative splicing difference  
A:Reference number: S06410; MUID:88302448  
A:Accession: S06907  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-492 <IZU>  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; nucleus; thyroid hormone receptor; transcription regulation  
F:51-327/Domain: erba transforming protein homology <ERBA>  
F:53-73/Region: zinc finger  
F:91-115/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 492;  
Best Local Similarity 83.3%; Pred. No. 13;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
|:|||||  
Db 110 CRFKKC 115

RESULT 12  
A40377  
thyroid hormone receptor beta-1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Dec-1991 #sequence\_revision 30-Dec-1991 #text\_change 29-Jan-1999

C:Accession: A40377  
R:Wood, W.M.; Ocran, K.W.; Gordon, D.F.; Ridgway, E.C.  
Mol. Endocrinol. 5, 1049-1061, 1991  
A:Title: Isolation and characterization of mouse complementary DNAs encoding alpha and beta thyroid hormone receptors from rat pituitary tumor cells.  
A:Reference number: A40377; MUID:92049384  
A:Accession: A40377  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-461 <WOO>  
A:Cross-references: GB:S62756; NID:g238331; PID:g238332  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: DNA binding; nucleus; thyroid hormone receptor; transcription regulation; zinc finger  
F:105-381/Domain: erba transforming protein homology <ERBA>  
F:107-127/Region: zinc finger  
F:145-169/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 461;  
Best Local Similarity 83.3%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
|:|||||  
Db 164 CRFKKC 169

RESULT 13  
D36067  
thyroid hormone receptor beta-B - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 08-Sep-1997  
C:Accession: D36067  
R:Yaota, Y.; Shi, Y.; Brown, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7090-7094, 1990  
A:Title: Xenopus laevis alpha and beta thyroid hormone receptors.  
A:Reference number: A36067; MUID:90384953  
A:Accession: D36067  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-373 <YAO>  
A:Cross-references: GB:M35361; NID:g213978; PID:g213979  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; nucleus; thyroid hormone receptor; transcription regulation  
F:17-293/Domain: erba transforming protein homology <ERBA>  
F:19-39/Region: zinc finger  
F:57-81/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 373;  
Best Local Similarity 83.3%; Pred. No. 10;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
|:|||||  
Db 76 CRFKKC 81

RESULT 14  
C36067  
thyroid hormone receptor beta-A - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 08-Mar-1991 #sequence\_revision 30-Sep-1993 #text\_change 16-Feb-1997  
C:Accession: C36067  
R:Yaota, Y.; Shi, Y.; Brown, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7090-7094, 1990  
A:Title: Xenopus laevis alpha and beta thyroid hormone receptors.  
A:Reference number: A36067; MUID:90384953  
A:Accession: C36067  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-373 <YAO>  
A:Cross-references: GB:M35359

C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; nucleus; thyroid hormone receptor; transd  
F:17-293/Domain: erba transforming protein homology <ERBA>  
F:19-39/Region: zinc finger  
F:57-81/Region: zinc finger

Query Match 92.3%; Score 36; DB 2; Length 373;  
Best Local Similarity 83.3%; Pred. No. 10;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
I:||||  
Db 76 CRFKKC 81

## RESULT 15

S35744  
Transforming protein (erba) - avian erythroblastosis virus  
C:Species: avian erythroblastosis virus  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 31-Oct-1997  
C:Accession: S35744  
R:Vennstroem, B.  
submitted to the EMBL Data Library, March 1993  
A:Reference number: S35743  
A:Accession: S35744  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-387 <VEN>  
A:Cross-references: EMBL:X12707  
C:Superfamily: thyroid hormone receptor; erba transforming protein homology  
C:Keywords: zinc finger  
F:37-313/Domain: erba transforming protein homology <ERBA>

Query Match 92.3%; Score 36; DB 2; Length 387;  
Best Local Similarity 83.3%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
I:||||  
Db 96 CRFKKC 101

Search completed: September 7, 1999, 23:06:28  
Job time: 2476 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:16 ; Search time 71.87 Seconds  
(without alignments)  
2.360 Million cell updates/sec

Title: US-09-124-280A-14  
Perfect score: 39  
Sequence: 1 CKFKKC 6

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	92.3	534	1 CNRD_CAEEL	P41830 caenorhabdi
2	36	92.3	864	1 E78A_DROME	P45447 drosophila
3	36	92.3	614	1 EAR1_HUMAN	P20393 homo sapien
4	36	92.3	385	1 ERBA_AVIER	P03373 avian erythr
5	36	92.3	499	1 FTF_MOUSE	P45448 mus musculu
6	36	92.3	410	1 THAI_HUMAN	P21205 homo sapien
7	36	92.3	410	1 THAI_MOUSE	P16416 mus musculu
8	36	92.3	410	1 THAI_SHEEP	Q28570 ovine aries
9	36	92.3	410	1 THA2_HUMAN	P10827 homo sapien
10	36	92.3	452	1 THA2_MOUSE	P37241 mus musculu
11	36	92.3	492	1 THA2_RAT	P15827 rattus norv
12	36	92.3	418	1 THAA_XENLA	P15204 xenopus lae
13	36	92.3	418	1 THAB_XENLA	P18116 xenopus lae
14	36	92.3	408	1 THA_CHICK	P04625 gallus gall
15	36	92.3	418	1 THA_RANCA	Q02777 rana catesb
16	36	92.3	461	1 THB1_HUMAN	P10828 homo sapien
17	36	92.3	461	1 THB1_MOUSE	P37242 mus musculu
18	36	92.3	461	1 THB1_RAT	P18113 rattus norv
19	36	92.3	411	1 THB1_SHEEP	Q28571 ovine aries
20	36	92.3	369	1 THB1_XENLA	P18116 xenopus lae
21	36	92.3	476	1 THB2_HUMAN	P37243 homo sapien
22	36	92.3	475	1 THB2_MOUSE	P37244 mus musculu
23	36	92.3	514	1 THB2_RAT	P37826 rattus norv
24	36	92.3	373	1 THB5_XENLA	P18117 xenopus lae
25	36	92.3	373	1 THB6_XENLA	P18118 xenopus lae
26	36	92.3	414	1 THB7_XENLA	P18119 xenopus lae
27	36	92.3	369	1 THB_CHICK	P18112 gallus gall
28	36	92.3	373	1 THB_RANCA	Q02965 rana catesb
29	35	89.7	4753	1 LRP_CAEEL	Q04833 caenorhabdi
30	34	87.2	1178	1 TSPI_CHICK	P35440 gallus gall
31	33	84.6	1257	1 FLIH_CAEEL	P34268 caenorhabdi
32	33	84.6	758	1 GCR_ONCMV	P49843 oncorhynch
33	33	84.6	310	1 YRRI_EBV	P03207 epstein-bar
34	32	82.1	461	1 AD4B_BOVIN	Q04752 bos taurus
35	32	82.1	461	1 AD4B_HUMAN	Q13285 homo sapien
36	32	82.1	462	1 AD4B_MOUSE	P33242 m sterold h
37	32	82.1	462	1 AD4B_RAT	P50569 rattus norv
38	32	82.1	225	1 FGFH_HUMAN	Q92914 homo sapien
39	32	82.1	225	1 FGFH_MOUSE	P70378 mus musculu
40	32	82.1	555	1 FTF1_BONMO	P49867 bombyx mori
41	32	82.1	1043	1 FTF1_DROME	P33244 drosophila
42	32	82.1	808	1 FTFB_DROME	Q05192 drosophila
43	32	82.1	501	1 FTF_CHICK	Q42101 gallus gall

ALIGNMENTS

RESULT 1			
ID	CNRD_CAEEL	STANDARD:	PRT: 534 AA.
AC	P41830;		
DT	01-NOV-1995 (REL. 32, CREATED)		
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)		
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)		
DE	STERIOD HORMONE RECEPTOR FAMILY MEMBER CNR14.		
GN	CNR-14 OR SEX-1 OR F44A6.2.		
OS	CAENORHABDITIS ELEGANS.		
OC	EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;		
OC	RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RX	MEDLINE: 95116514.		
RA	KOSTROUCH Z., KOSTROUCHOVA M., RALL J.E.;		
RL	"Steroid/thyroid hormone receptor genes in Caenorhabditis elegans.";		
RL	PROC. NATL. ACAD. SCI. U.S.A. 92:156-159(1995).		
RN	[2]		
RP	REVISIONS TO C-TERMINAL.		
RA	KOSTROUCHOVA M.;		
RL	SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	SULSTON J.;		
RL	SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.		
CC	- - SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).		
CC	- - TISSUE SPECIFICITY: MOST ABUNDANT IN EMBRYOS.		
CC	- - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U13074; G1203792; -		
DR	EMBL; Z50858; E1346642; -		
DR	WORMPEP; F44A6.2; CE03323.		
DR	PROSITE; P500031; NUCLEAR_RECEPTOR; 1.		
DR	PFAM; PF00104; hormone_rec; 1.		
DR	PFAM; PF00105; zf-C4; 1.		
DR	HSP; P03372; LHCO.		
KW	RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;		
KW	ZINC-FINGER.		
FT	DNA_BIND 151 216 C4-TYPE ZINC FINGERS (TWO).		
FT	ZN_FING 151 171 C4-TYPE.		
FT	ZN_FING 187 211 C4-TYPE.		
SQ	SEQUENCE 534 AA; 60848 MW; BF597178 CRC32;		
Query Match 92.3%; Score 36; DB 1; Length 534;			
Best Local Similarity 83.3%; Pred. No. 9.7;			
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
QY	1 CKFKKC 6		
	:		
Db	206 CRFKKC 211		
RESULT 2			
E78A_DROME			

ID E78A\_DROME STANDARD; PRT; 864 AA.  
 AC P45447;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE ECDYSONE-INDUCIBLE PROTEIN E78-A (DR-78).  
 GN EIP78C OR E78A.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S;  
 RX MEDLINE; 94006562.  
 RA STONE B.L., THUMMEL C.S.;  
 RT "The Drosophila 78C early late puff contains E78, an  
 RT ecdysone-inducible gene that encodes a novel member of the nuclear  
 RT hormone receptor superfamily.";  
 RL CELL 75:307-320(1993).  
 RN [2]  
 RP SEQUENCE OF 321-433 FROM N.A.  
 RC MEDLINE; 94060116.  
 RA MARTIN-BLANCO E., KORNBERG T.B.;  
 RT "DR-78, a novel Drosophila melanogaster genomic DNA fragment highly  
 RT homologous to the DNA-binding domain of thyroid hormone-retinoic  
 RT acid-vitamin D receptor subfamily.";  
 RL BIOCHIM. BIOPHYS. ACTA 1216:339-341(1993).  
 CC -1- FUNCTION: INDUCES THE EARLY LATE PUFF 78C WHICH TRIGGERS PUPARIUM  
 CC FORMATION AND DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -1- DEVELOPMENTAL STAGE: THE LONGER FORM, E78A, IS EXPRESSED ONLY  
 CC IN MID-PUPAL STAGES, WHILEST THE SHORTER FORM, E78B, IS MAXIMALLY  
 CC EXPRESSED IN NEWLY FORMED PREPUPAE.  
 CC -1- INDUCTION: BOTH FORMS REQUIRE ECDYSONE FOR ACTIVITY. E78B ALSO  
 CC REQUIRES ECDYSONE-INDUCED PROTEINS FOR MAXIMAL EXPRESSION.  
 CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT FORMS OF PROTEIN E78 ARE  
 CC PROBABLY PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U01087; G514833; -;  
 DR EMBL; U01088; G514834; -;  
 DR EMBL; X73045; E74981; -;  
 DR FLYBASE; FBgn0004865; Eip78C.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P03372; IHCP.  
 KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; ALTERNATIVE SPLICING.  
 FT DNA\_BIND 367 432 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 367 387 C4-TYPE.  
 FT ZN\_FING 403 427 C4-TYPE.  
 FT DOMAIN 64 80 POLY-GLU.  
 FT DOMAIN 182 188 POLY-GLN.  
 FT DOMAIN 192 202 POLY-GLN.  
 FT DOMAIN 240 247 POLY-SER.  
 FT DOMAIN 271 279 POLY-SER.  
 FT DOMAIN 312 315 POLY-GLN.  
 FT DOMAIN 321 333 POLY-GLN.  
 FT DOMAIN 336 339 POLY-GLN.  
 FT DOMAIN 346 349 POLY-SER.  
 FT DOMAIN 354 357 POLY-ASN.  
 FT DOMAIN 481 486 POLY-GLN.  
 FT DOMAIN 490 506 POLY-GLN.

FT DOMAIN 546 554 POLY-ASN.  
 FT VARSPLIC 1 474 MISSING (IN TRUNCATED FORM E78B).  
 FT CONFLICT 321 331 QLOQOOHQOO -> SCNSSSTSSR (IN REF. 2).  
 FT CONFLICT 430 433 AGMS -> YGKM (IN REF. 2).  
 SQ SEQUENCE 864 AA; 95865 MW; 5EBE72C0 CRC32;  
  
 Query Match 92.3%; Score 36; DB 1; Length 864;  
 Best Local Similarity 83.3%; Pred. No. 14;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 CKFKKC 6  
 DB 422 CKFKKC 427  
  
 RESULT 3  
 EARL\_HUMAN STANDARD; PRT; 614 AA.  
 ID EARL\_HUMAN AC P20393;  
 DT 01-FEB-1991 (REL. 17, CREATED)  
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE V-ERBA RELATED PROTEIN EAR-1.  
 GN EAR1 OR HREV.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE; 89195219.  
 RA MIYAJIMA N., HORIUCHI R., SHIBUYA Y., FUKUSHIGE S.-I.,  
 RA MATSUBARA K.-I., TOYOSHIMA K., YAMAMOTO T.;  
 RT "Two erba homologs encoding proteins with different T3 binding  
 RT capacities are transcribed from opposite DNA strands of the same  
 RT genetic locus.";  
 RL CELL 57:31-39(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKELETAL MUSCLE;  
 RX MEDLINE; 90262650.  
 RA LAZAR M.A., JONES K.E., CHIN W.W.;  
 RT "Isolation of a cDNA encoding human Rev-Erba alpha: transcription  
 RT from the noncoding DNA strand of a thyroid hormone receptor gene  
 RT results in a related protein that does not bind thyroid hormone.";  
 RL DNA CELL BIOL. 9:77-83(1990).  
 RN [3]  
 RP SEQUENCE OF 549-614 FROM N.A.  
 RX MEDLINE; 91212192.  
 RA LAUDET V., BEGUE A., HENRY C., JOUBEL A., MARTIN P., STEHELIN D.,  
 RA SAULE S.;  
 RT "Genomic organization of the human thyroid hormone receptor alpha (c-  
 RT erba-1) gene.";  
 RL NUCLEIC ACIDS RES. 19:1105-1112(1991).  
 CC -1- FUNCTION: PROBABLE RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M24898; G537520; -;  
 DR EMBL; M24900; G537516; -;  
 DR EMBL; X55086; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; X72631; -; NOT\_ANNOTATED\_CDS.

DR EMBL; X72632; G732803; ALT\_SEQ.  
DR MIM; A32286; A32286.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR PFAM; PF00104; hormone\_rec; 1.  
DR PFAM; PF00105; zf-C4; 1.  
DR HSSP; P03372; LHCO.  
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
ZINC-FINGER.  
FT DOMAIN 82 93 POLY-SER.  
FT DNA\_BIND 132 198 C4-TYPE ZINC FINGERS (TWO).  
FT ZN\_FING 132 152 C4-TYPE.  
FT ZN\_FING 169 193 C4-TYPE.  
FT ZN\_FING 147 177 H -> L (IN REF. 2).  
FT CONFLICT 147 177 E -> Q (IN REF. 2).  
FT CONFLICT 564 564 E -> Q (IN REF. 2).  
SQ SEQUENCE 614 AA; 66805 MW; 4D5A21EE CRC32;  
  
Query Match 92.3%; Score 36; DB 1; Length 614;  
Best Local Similarity 83.3%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CRFKKC 6  
Db 188 CRFKKC 193  
  
RESULT 4  
ERBA\_AVIER STANDARD; PRT; 385 AA.  
AC P03373;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE ERBA ONCOGENE PROTEIN.  
GN V-ERBA.  
OS AVIAN ERYTHROBLASTOSIS VIRUS (STRAIN ES4).  
OC VIRUSES; RETROVIRUSES; RETROVIRIDAE; AVIAN TYPE C RETROVIRUSES.  
RN [1]  
RX MEDLINE; 87218491.  
RA DAMM K., BEUG H., GRAF T., VENNSTROEM B.;  
RT "A single point mutation in erba restores the erythroid transforming  
potential of a mutant avian erythroblastosis virus (AEV) defective in  
both erba and erbb oncogenes.";  
RL EMBO J. 6:375-382(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 84223957.  
RA DEBUIRE B., HENRY C., BENAÏSSA M., BISERTE G., CLAVERIE J.-M.,  
RA SAULE S., MARTIN P., STEHELIN D.;  
RT "Sequencing the erba gene of avian erythroblastosis virus reveals a  
new type of oncogene.";  
RL SCIENCE 224:1456-1459(1984).  
RN [3]  
RP ACTION.  
RX MEDLINE; 89281730.  
RA DAMM K., THOMPSON C.C., EVANS R.M.;  
RT "Protein encoded by v-erba functions as a thyroid-hormone receptor  
antagonist.";  
RL NATURE 339:593-597(1989).  
CC -!- FUNCTION: V-ERBA ACTS AS A THYROID-HORMONE RECEPTOR ANTAGONIST.  
CC -!- THE V-ERBA ONCOGENE COOPERATES WITH V-ERBB AND OTHER PRIMARILY  
CC SARCOMA-INDUCING ONCOGENES IN TRANSFORMATION OF ERYTHROBLASTS  
CC V-ERBA IS NON TRANSFORMING BY ITSELF AND CANNOT INDUCE SELF-  
CC RENEWAL OF ERYTHROBLASTS. INSTEAD IT BLOCKS THE RESIDUAL  
CC CAPACITY OF THE V-ERBB CELLS TO DIFFERENTIATE TERMINALLY.  
CC -!- C-ERBA, THE CELLULAR COUNTERPART OF V-ERBA, IS A THYROID HORMONE  
CC RECEPTOR.  
CC -!- THIS PROTEIN IS SYNTHESIZED AS A GAG-ERBA-ERBB POLYPEPTIDE.  
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 381  
CC ONWARD AND IS LONGER (398 AA) DUE TO A FRAMESHIFT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; Y00044; G871521; ALT\_INIT.  
DR PIR; A03248; TFEVVR.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR PFAM; PF00104; hormone\_rec; 1.  
DR PFAM; PF00105; zf-C4; 1.  
DR HSSP; P10828; 2NLL.  
KW RECEPTOR; TRANSCRIPTION REGULATION; ONCOGENE; THYROID HORMONE; REPRESSOR;  
ZINC-FINGER; POLYPEPTIDE; DNA-BINDING; NUCLEAR PROTEIN;  
FT DNA\_BIND 37 104 C4-TYPE ZINC FINGERS (TWO).  
FT ZN\_FING 37 57 C4-TYPE.  
FT ZN\_FING 75 99 C4-TYPE.  
SQ SEQUENCE 385 AA; 44029 MW; 2BAAB074 CRC32;  
  
Query Match 92.3%; Score 36; DB 1; Length 385;  
Best Local Similarity 83.3%; Pred. No. 7.5;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CRFKKC 6  
Db 94 CRFKKC 99  
  
RESULT 5  
FTF\_MOUSE STANDARD; PRT; 499 AA.  
ID FTF\_MOUSE  
AC P45448;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE ALPHA-1-FETOPROTEIN TRANSCRIPTION FACTOR (LIVER RECEPTOR HOMOLOG)  
(LRH-1).  
GN FTF.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TUGWOOD J.D., ISEEMANN I., GREEN S.;  
RL SUBMITTED (XXX-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
CC FTF-F1 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; M81385; G198875; ALT\_INIT.  
DR MGD; MGI:1203501; FTF.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
DR PFAM; PF00104; hormone\_rec; 1.  
DR PFAM; PF00105; zf-C4; 1.  
DR HSSP; P19793; 2NLL.  
KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
ZINC-FINGER.  
FT DNA\_BIND 46 111 C4-TYPE ZINC FINGERS (TWO).  
FT ZN\_FING 46 66 C4-TYPE.  
FT ZN\_FING 82 106 C4-TYPE.  
SQ SEQUENCE 499 AA; 57175 MW; 01832061 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 499;  
 Best Local Similarity 83.3%; Pred. No. 9.2;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 101 CKFKKC 106

RESULT 6  
 ID THAI\_HUMAN STANDARD; PRT; 410 AA.  
 AC P21205;  
 DT 01-MAY-1991 (REL. 18, CREATED)  
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-1 (EAR-7-1) (C-ERBA-ALPHA-1).  
 GN THRA OR THRAI OR ERBAL.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91212192.  
 RA LAUDET V., BEGUE A., HENRY C., JOUBEL A., MARTIN P., STEHELIN D.,  
 RA SAULE S.;  
 RT "Genomic organization of the human thyroid hormone receptor alpha (c-  
 erba-1) gene.";  
 RL NUCLEIC ACIDS RES. 19:1105-1112(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89195219.  
 RA MIYAJIMA N., HORIUCHI R., SHIBUYA Y., FUKUSHIGE S.-I.,  
 RA MATSUBARA K.-I., TOYOSHIMA K., YAMAMOTO T.;  
 RT "Two erba homologs encoding proteins with different T3 binding  
 capacities are transcribed from opposite DNA strands of the same  
 genetic locus.";  
 RL CELL 57:31-39(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89127255.  
 RA NAKAI A., SAKURAI A., BELL G.I., DEGROOT L.J.;  
 RT "Characterization of a third human thyroid hormone receptor  
 coexpressed with other thyroid hormone receptors in several  
 tissues.";  
 RL MOL. ENDOCRINOL. 2:1087-1092(1988).  
 RN [4]  
 RP SEQUENCE OF 1-370 FROM N.A.  
 RC TISSUE=KIDNEY;  
 RX MEDLINE; 88190152.  
 RA NAKAI A., SAKURAI A., SZILAK I., BELL G.I., DEGROOT L.J.;  
 RT "Characterization of a thyroid hormone receptor expressed in human  
 kidney and other tissues.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 85:2781-2785(1988).  
 RN [5]  
 RP SEQUENCE OF 1-370 FROM N.A.  
 RC TISSUE=TESTIS;  
 RX MEDLINE; 88067793.  
 RA PFAHL M., BENBROOK D.;  
 RT "Nucleotide sequence of cDNA encoding a novel human thyroid hormone  
 receptor.";  
 RL NUCLEIC ACIDS RES. 15:9613-9613(1987).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- ALTERNATIVE PRODUCTS: ALPHA-1 AND ALPHA-2 RECEPTORS ARE PRODUCED  
 CC BY ALTERNATIVE SPLICING OF THE SAME GENE.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X55005; G29879; -  
 DR EMBL; X55074; G825639; -  
 DR EMBL; X55073; G825639; JOINED.  
 DR EMBL; X55070; G825639; JOINED.  
 DR EMBL; X55071; G825639; JOINED.  
 DR EMBL; X55004; G825639; JOINED.  
 DR EMBL; X55069; G825639; JOINED.  
 DR EMBL; X55068; G825639; JOINED.  
 DR EMBL; M24748; G339663; -  
 DR EMBL; Y00479; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A30893; A30893.  
 DR PIR; A40917; A40917.  
 DR PIR; S06247; S06247.  
 DR PIR; B32286; B32286.  
 DR PIR; S26828; S26828.  
 DR MIM; 190120; -  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSSP; P10828; 2NLL.  
 DR TRANSFAC; T00838; -  
 DR TRANSFAC; T01152; -  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 FT ZINC-FINGER; ALTERNATIVE SPLICING; MULTIGENE FAMILY.  
 FT DOMAIN 1 52  
 FT DNA\_BIND 53 120  
 FT ZN\_FING 53 73  
 FT ZN\_FING 91 115  
 FT DOMAIN 190 370  
 FT CONFLICT 285 285  
 FT CONFLICT E -> A (IN REF. 1).  
 SQ SEQUENCE 410 AA; 46813 MW; 7C1D1840 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 7.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 110 CKFKKC 115

RESULT 7  
 ID THAI\_MOUSE STANDARD; PRT; 410 AA.  
 AC P16416; P10685; O63107;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-1.  
 GN THRA OR C-ERBA-ALPHA.  
 OS MUS MUSCULUS (MOUSE); AND RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=MOUSE; STRAIN=C3H; TISSUE=TESTIS;  
 RX MEDLINE; 90272421.  
 RA MASUDA M., YASUHARA S., YAMASHITA M., SHIBUYA M., ODAKA T.;  
 RT "Nucleotide sequence of the murine thyroid hormone receptor (alpha-1)  
 cDNA.";  
 RL NUCLEIC ACIDS RES. 18:3055-3055(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=MOUSE; STRAIN=C3H; TISSUE=MUSCLE;  
 RX MEDLINE; 88289388.

RA PROST E., KOENIG R.J., MOORE D.D., LARSEN P.R., WHALEN R.G.;  
 RT "Multiple sequences encoding potential thyroid hormone receptors  
 RT isolated from mouse skeletal muscle cDNA libraries.";  
 RL NUCLEIC ACIDS RES. 16:6248-6248(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-RAT;  
 RX MEDLINE: 88302448.  
 RA IZUMO S., MAHDAMI V.;  
 RT "Thyroid hormone receptor alpha isoforms generated by alternative  
 RT splicing differentially activate myosin HC gene transcription.";  
 RL NATURE 334:539-542(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-RAT;  
 RX MEDLINE: 87319689.  
 RA THOMPSON C.C., WEINBERGER C., LEBO R., EVANS R.M.;  
 RT "Identification of a novel thyroid hormone receptor expressed in the  
 RT mammalian central nervous system.";  
 RL SCIENCE 237:1610-1614(1987).  
 RN [5]  
 RP SEQUENCE OF 1-370 FROM N.A.  
 RC SPECIES-RAT;  
 RX MEDLINE: 88262581.  
 RA MITSUHASHI T., TENNYSON G., NIKODEM V.;  
 RT "Nucleotide sequence of novel cDNAs generated by alternative splicing  
 RT of a rat thyroid hormone receptor gene transcript.";  
 RL NUCLEIC ACIDS RES. 16:5697-5697(1988).  
 RN [6]  
 RP SEQUENCE OF 13-410 FROM N.A.  
 RC SPECIES-RAT; TISSUE=HEART;  
 RA FLINK I.L., BAILEY T., BAHJ J., MORKIN E.;  
 RT SUBMITTED (AUG-1988) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 152-410.  
 RC SPECIES-RAT;  
 RX MEDLINE: 96107308.  
 RA WAGNER R.L., APRILETTI J.W., MCGRATH M.E., WEST B.L., BAXTER J.D.,  
 RA FLETTERICK R.J.;  
 RT "A structural role for hormone in the thyroid hormone receptor.";  
 RL NATURE 378:690-697(1995).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- ALTERNATIVE PRODUCTS: ALPHA-1, ALPHA-2 AND ALPHA-3 RECEPTORS ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING. THE THREE SEQUENCES ARE  
 CC IDENTICAL UP TO POSITION 370.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X51983; G50386; -;  
 DR EMBL: X07750; G50388; -;  
 DR EMBL: M18028; G207301; -;  
 DR EMBL: X07409; G57391; ALT\_TERM.  
 DR EMBL: X12744; G55932; -;  
 DR PIR: S14690; ORMSA1  
 DR PIR: S14416; S14416.  
 DR PIR: S06410; S06410.  
 DR MGD; MGI:98742; THRA.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSSP; P10828; 2NLL.  
 DR TRANSFAC; T01173; -;

KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; ALTERNATIVE SPLICING; MULTIGENE FAMILY.  
 FT DOMAIN 1 52  
 FT DNA\_BIND 53 120  
 FT ZN\_FING 53 73  
 FT ZN\_FING 91 115  
 FT ZN\_FING 190 410  
 FT DOMAIN 190 410  
 FT CONFLICT 106 106  
 FT CONFLICT 281 281  
 FT CONFLICT 285 285  
 FT CONFLICT 285 285  
 SQ SEQUENCE 410 AA; 46795 MW; F805F19B CRC32;  
 Query Match 92.3%; Score 36; DB 1; Length 410;  
 Best Local Similarity 83.3%; Pred. No. 7.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CRFKKC 6  
 Db 110 CRFKKC 115  
 RESULT 8  
 THAL\_SHEEP  
 ID THAL\_SHEEP STANDARD; PRT; 410 AA.  
 AC Q28570;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-1.  
 GN THRA OR C-ERBA-ALPHA.  
 OS OVIS ARIES (SHEEP).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; CAPRINAE; OVIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE: 96434366.  
 RA TUCKER M.A., POLK D.H.;  
 RT "Cloning of the alpha and beta ovine thyroid hormone receptor cDNAs.";  
 RL THYROID 6:237-243(1996).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- ALTERNATIVE PRODUCTS: ALPHA-1 AND ALPHA-2 RECEPTORS ARE PRODUCED  
 CC BY ALTERNATIVE SPLICING OF THE SAME GENE. THE TWO SEQUENCES ARE  
 CC IDENTICAL UP TO POSITION 370.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z68308; E214826; -;  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSSP; P10828; 2NLL.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; ALTERNATIVE SPLICING; MULTIGENE FAMILY.  
 FT DOMAIN 1 52  
 FT DNA\_BIND 53 120  
 FT ZN\_FING 53 73  
 FT ZN\_FING 91 115  
 FT ZN\_FING 190 410  
 FT DOMAIN 190 410  
 FT CONFLICT 106 106  
 FT CONFLICT 281 281  
 FT CONFLICT 285 285  
 SQ SEQUENCE 410 AA; 46802 MW; EEB4B7C1 CRC32;

```

Query Match          92.38; Score 36; DB 1; Length 410;
Best Local Similarity 83.3%; Pred. No. 7.9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CKPKKC 6
Db      110 CRFKKC 115
       1:1111
       1:1111

RESULT 9
THA2_HUMAN
ID THA2_HUMAN STANDARD; PRT; 490 AA.
AC P10827;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE THYROID HORMONE RECEPTOR ALPHA-2 (EAR-7-2) (C-ERBA-ALPHA-2).
GN THRA OR THRA1 OR ERBA1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 91212192.
RA LAUDET V., BEGUE A., HENRY C., JOUBEL A., MARTIN P., STEHELIN D.,
RA SAULE S.;
RN RP SEQUENCE FROM N.A.
RX MEDLINE; 89195219.
RA MIYAJIMA N., HORIUCHI R., SHIBUYA Y., FUKUSHIGE S.-I.,
RA MATSUBARA K.-I., TOYOSHIMA K., YAMAMOTO T.;
RT "Two erba homologs encoding proteins with different T3 binding
RT capacities are transcribed from opposite DNA strands of the same
RT genetic locus.";
RN RL CELL 57:31-39(1989).
[3]
RN RP SEQUENCE FROM N.A.
RX TISSUE-KIDNEY;
RX MEDLINE; 88190152.
RA NAKAI A., SEINO S., SAKURAI A., SZILAK I., BELL G.I., DEGROOT L.J.;
RT "Characterization of a thyroid hormone receptor expressed in human
RT kidney and other tissues.";
RN RL PROC. NATL. ACAD. SCI. U.S.A. 85:2781-2785(1988).
[4]
RN RP SEQUENCE FROM N.A.
RX TISSUE=TESTIS;
RA PFAHL M., BENBROOK D.;
RT "Nucleotide sequence of cDNA encoding a novel human thyroid hormone
RT receptor.";
RN RL NUCLEIC ACIDS RES. 15:9613-9613(1987).
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- ALTERNATIVE PRODUCTS: ALPHA-1 AND ALPHA-2 RECEPTORS ARE PRODUCED
CC BY ALTERNATIVE SPLICING OF THE SAME GENE.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M24899; G537522; -
DR EMBL; X55004; -; NOT_ANNOTATED_CDS.
DR EMBL; X55066; -; NOT_ANNOTATED_CDS.

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X07751; G50390; -.  
 DR PIR; S14417; S14417.  
 DR MGD; MGI:98742; THRA.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR HSPAM; PF00105; zf-C4; 1.  
 DR HSP; P10828; 2NLL.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; ALTERNATIVE SPLICING; MULTIGENE FAMILY.  
 FT DOMAIN 1 52  
 FT DNA\_BIND 53 120 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 53 73 C4-TYPE.  
 FT ZN\_FING 91 115 C4-TYPE.  
 FT DOMAIN 190 370 LIGAND-BINDING.  
 FT NON\_TER 452 452  
 SQ SEQUENCE 452 AA; 50981 MW; F00DC957 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 452;

Best Local Similarity 83.3%; Pred. No. 8.5; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CKFKKC 6  
 I:||||  
 Db 110 CRFKKC 115

# RESULT 11

THA2\_RAT  
 ID THA2\_RAT STANDARD; PRT; 492 AA.  
 AC P15827; O63195; Q63196;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-2.  
 GN THRA OR C-ERBA-ALPHA.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88302448.  
 RA IZUMO S., MAHDAVI V.;  
 RT "Thyroid hormone receptor alpha isoforms generated by alternative  
 RT splicing differentially activate myosin HC gene transcription.";  
 RL NATURE 334:539-542(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88262581.  
 RA MITSUHASHI T., TENNYSON G., NIKODEM V.;  
 RT "Nucleotide sequence of novel cDNAs generated by alternative splicing  
 RT of a rat thyroid hormone receptor gene transcript.";  
 RL NUCLEIC ACIDS RES. 16:5697-5697(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89039899.  
 RA LAZAR M.A., HODIN R.A., DARLING D.S., CHIN W.W.;  
 RT "Identification of a rat c-erba alpha-related protein which binds  
 RT deoxyribonucleic acid but does not bind thyroid hormone.";  
 RL MOL. ENDOCRINOL. 2:893-901(1988).  
 CC -|- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -|- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -|- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -|- ALTERNATIVE PRODUCTS: ALPHA-1 AND ALPHA-2 RECEPTORS ARE PRODUCED  
 CC BY ALTERNATIVE SPLICING OF THE SAME GENE. THE TWO SEQUENCES ARE  
 CC IDENTICAL UP TO POSITION 370.  
 CC -|- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X07409; G57391; ALT\_INIT.  
 DR EMBL; M31174; G204048; -.  
 DR EMBL; M31177; G204050; -.  
 DR PIR; S00927; TVRTAR.  
 DR PIR; S06907; S06907.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR HSPAM; PF00105; zf-C4; 1.  
 DR HSP; P10828; 2NLL.  
 DR TRANSFAC; T00840; -.  
 DR TRANSFAC; T00841; -.  
 DR TRANSFAC; T01343; -.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; ALTERNATIVE SPLICING; MULTIGENE FAMILY.  
 FT DOMAIN 1 52  
 FT DNA\_BIND 53 120 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 53 73 C4-TYPE.  
 FT ZN\_FING 91 115 C4-TYPE.  
 FT DOMAIN 190 370 LIGAND-BINDING.  
 FT CONFLICT 72 72 G -> C (IN REF. 3).  
 FT CONFLICT 285 285 E -> K (IN REF. 1).  
 FT CONFLICT 482 482 F -> C (IN REF. 3).  
 SQ SEQUENCE 492 AA; 55071 MW; 32094103 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 492;

Best Local Similarity 83.3%; Pred. No. 9.1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFKKC 6

I:||||

Db 110 CRFKKC 115

# RESULT 12

THAA\_XENLA  
 ID THAA\_XENLA STANDARD; PRT; 418 AA.  
 AC P15204; P18114;  
 DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-A.  
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90384953.  
 RA YAOITA Y., SHI Y.-B., BROWN D.D.;  
 RT "Xenopus laevis alpha and beta thyroid hormone receptors.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:7090-7094(1990).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE; 91046059.  
 RA YAOITA Y., SHI Y.-B., BROWN D.D.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:8684-8684(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OOCYTE;  
 RX BROOKS A.R., SWENEY G., OLD R.W.;  
 RT "Structure and functional expression of a cloned Xenopus thyroid  
 RT hormone receptor.";  
 RL NUCLEIC ACIDS RES. 17:9395-9405(1989).  
 CC -|- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -|- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M35343; G214830; -.  
 CC PIR: X17385; G65142; -.  
 CC PIR: S09605; TVXLT.A.  
 CC PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 CC PFAM: PF00104; hormone\_rec; 1.  
 CC HSP: P10828; 2NL.  
 CC RECEPTOR: TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 CC ZINC-FINGER; MULTIGENE FAMILY.  
 CC DOMAIN 1 60 MODULATING.  
 CC DNA\_BIND 61 128 C4-TYPE ZINC FINGERS (TWO).  
 CC ZN\_FING 61 81 C4-TYPE.  
 CC ZN\_FING 99 123 C4-TYPE.  
 CC DOMAIN 129 418 LIGAND-BINDING.  
 CC CONFLICT 54 54 Y -> H (IN REF. 3).  
 CC CONFLICT 137 137 G -> S (IN REF. 3).  
 CC SEQUENCE 418 AA; 47672 MW; 3FB563BC CRC32;

Query Match 92.3%; Score 36; DB 1; Length 418;  
 Best Local Similarity 83.3%; Pred. No. 8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 DB 118 CRFKKC 123  
 :|:|:|

RESULT 13  
 THAB\_XENLA  
 ID THAB\_XENLA STANDARD; PRT; 418 AA.  
 AC P18115;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA-B.  
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90384953.  
 RA YAOITA Y., SHI Y.-B., BROWN D.D.;  
 RT "Xenopus laevis alpha and beta thyroid hormone receptors.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:7090-7094(1990).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE; 91046059.  
 RA YAOITA Y., SHI Y.-B., BROWN D.D.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:8684-8684(1990).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M35344; G214832; -.  
 CC PIR: B36067; B36067.  
 CC PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 CC PFAM: PF00104; hormone\_rec; 1.  
 CC HSP: P10828; 2NL.  
 CC RECEPTOR: TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 CC ZINC-FINGER; MULTIGENE FAMILY.  
 CC DOMAIN 1 60 MODULATING.  
 CC DNA\_BIND 61 128 C4-TYPE ZINC FINGERS (TWO).  
 CC ZN\_FING 61 81 C4-TYPE.  
 CC ZN\_FING 99 123 C4-TYPE.  
 CC DOMAIN 129 418 LIGAND-BINDING.  
 CC SEQUENCE 418 AA; 47594 MW; 85209B56 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 418;  
 Best Local Similarity 83.3%; Pred. No. 8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 DB 118 CRFKKC 123  
 :|:|:|

RESULT 14  
 THA\_CHICK  
 ID THA\_CHICK STANDARD; PRT; 408 AA.  
 AC P04625;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA.  
 GN C-ERBA-ALPHA.  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPAFAS;  
 RX MEDLINE; 87090374.  
 RA SAP J., MUNOZ A., DAMM K., GHYSDAEL J., LEUTZ A., BEUG H.,  
 RA VENNSTROEM B.;  
 RT "The c-erb-A protein is a high-affinity receptor for thyroid  
 RT hormone.";  
 RL NATURE 324:635-640(1986).  
 RN [2]  
 RP SEQUENCE OF 1-239 FROM N.A.  
 RX MEDLINE; 87246693.  
 RA ZAHRAOUI A., CUNY G.;  
 RT "Nucleotide sequence of the chicken proto-oncogene c-erbA  
 RT corresponding to domain 1 of v-erbA.";  
 RL EUR. J. BIOCHEM. 166:63-69(1987).  
 RN [3]  
 RP PHOSPHORYLATION  
 RX MEDLINE; 90016293.  
 RA GLINEUR C., BAILLY M., GHYSDAEL J.;  
 RT "The c-erbA alpha-encoded thyroid hormone receptor is phosphorylated  
 RT in its amino terminal domain by casein kinase II.";  
 RL ONCOGENE 4:1247-1254(1989).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X04854; G63176; -;  
 DR EMBL; Y00987; G63178; -;  
 DR PIR; A25236; TVCHVR. -;  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P10828; 2NLL.  
 DR TRANSFAC; T01351; -;  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; MULTIGENE FAMILY; PROTO-ONCOGENE; PHOSPHORYLATION.  
 FT DOMAIN 1 50 MODULATING.  
 FT DNA\_BIND 51 118 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 51 71 C4-TYPE.  
 FT ZN\_FING 89 113 C4-TYPE.  
 FT DOMAIN 168 408 LIGAND-BINDING.  
 FT MOD\_RES 12 12 PHOSPHORYLATION (BY CK2).  
 FT MOD\_RES 28 28 PHOSPHORYLATION.  
 SQ SEQUENCE 408 AA; 46757 MW; B7C7E2ED CRC32;

Query Match 92.3%; Score 36; DB 1; Length 408;  
 Best Local Similarity 83.3%; Pred. No. 7.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
 I:|||||  
 DB 108 CRFKKC 113

RESULT 15  
 THA\_RANCA STANDARD; PRT; 418 AA.  
 AC Q02777;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE THYROID HORMONE RECEPTOR ALPHA.  
 OS RANA CATESBEIANA (BULL FROG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 OC NEOBATRACHIA; RANOIDEA; RANIDAE; RANINAE; RANA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-TADPOLE TAIL;  
 RX MEDLINE; 94062613.  
 RA SCHNEIDER M.J., DAVEY J.C., GALTON V.A.;  
 RT "Rana catesbeiana tadpole red blood cells express an alpha, but not a  
 RT beta, c-erba gene.";  
 RL ENDOCRINOLOGY 133:2488-2495(1993).  
 RN [2]  
 RP SEQUENCE OF 77-418 FROM N.A.  
 RC TISSUE-ERYTHROCYTE;  
 RX MEDLINE; 91246218.  
 RA SCHNEIDER M.J., GALTON V.A.;  
 RT "Regulation of c-erba-alpha messenger RNA species in tadpole  
 RT erythrocytes by thyroid hormone.";  
 RL MOL. ENDOCRINOL. 5:201-208(1991).  
 CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- TISSUE SPECIFICITY: HIGHEST LEVEL OF EXPRESSION IN ERYTHROCYTES.  
 CC ALSO EXPRESSED IN LIVER, TAIL, EYE, MUSCLE AND SKIN.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS LOW AT STAGE XII OF TADPOLE  
 CC DEVELOPMENT, GREATLY INCREASES BY STAGE XIX, DECREASES THEREAFTER  
 CC AND IS RELATIVELY LOW BY STAGE XXIV, WHEN METAMORPHOSIS IS  
 CC ALMOST COMPLETE.  
 CC -!- INDUCTION: BY THYROID HORMONE.  
 CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; L06064; G213654; -;  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P10828; 2NLL.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER; MULTIGENE FAMILY.  
 FT DOMAIN 1 60 MODULATING (BY SIMILARITY).  
 FT DNA\_BIND 61 128 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 61 81 C4-TYPE.  
 FT ZN\_FING 99 123 C4-TYPE.  
 FT DOMAIN 178 418 LIGAND-BINDING (BY SIMILARITY).  
 SQ SEQUENCE 418 AA; 47748 MW; 5C11BC99 CRC32;

Query Match 92.3%; Score 36; DB 1; Length 418;  
 Best Local Similarity 83.3%; Pred. No. 8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRFKKC 6  
 I:|||||  
 DB 118 CRFKKC 123

Search completed: September 7, 1999, 23:50:17  
 Job time: 1956 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:45 ; Search time 148.39 Seconds  
(without alignments)  
2.488 Million cell updates/sec

Title: US-09-124-280A-14

Perfect score: 39

Sequence: 1 CKFKKC 6

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTRMBL10:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	92.3	456	4	Q13986 homo sapien
2	36	92.3	579	4	O00402 homo sapien
3	36	92.3	579	4	Q14995 homo sapien
4	36	92.3	445	4	O95779 homo sapien
5	36	92.3	498	4	O95780 homo sapien
6	36	92.3	865	5	O18395 drosophila
7	36	92.3	810	5	O61228 tenebrio mo
8	36	92.3	389	5	O17929 caenorhabdi
9	36	92.3	556	5	O17013 caenorhabdi
10	36	92.3	363	5	O18609 caenorhabdi
11	36	92.3	422	5	O45460 caenorhabdi
12	36	92.3	338	5	O17928 caenorhabdi
13	36	92.3	342	5	O46003 caenorhabdi
14	36	92.3	452	5	O45315 caenorhabdi
15	36	92.3	1464	5	O98836 drosophila
16	36	92.3	410	6	O97715 sus scrofa
17	36	92.3	506	6	O97716 sus scrofa
18	36	92.3	494	11	O62702 rattus norv
19	36	92.3	508	11	O63503 rattus norv
20	36	92.3	212	11	O63620 rattus norv
21	36	92.3	578	11	O63504 rattus norv
22	36	92.3	383	11	O63542 rattus norv
23	36	92.3	413	11	O99146 mus musculu
24	36	92.3	533	11	O62756 rattus norv
25	36	92.3	576	11	O60674 mus musculu
26	36	92.3	962	12	O64895 avian eryth
27	36	92.3	455	12	O85511 avian eryth
28	36	92.3	582	12	O96593 avian eryth
29	36	92.3	555	12	O96594 avian eryth

#### ALIGNMENTS

```

RESULT 1
Q13986 PRELIMINARY; PRT: 456 AA.
AC Q13986
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE C-ERBA.
DE C-ERBA.
GN C-ERBA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87217099.
RA WEINBERGER C., GIGUERE V., HOLLENBERG S., ROSENFELD M.G., EVANS R.M.;
RT "Human steroid receptors and erba proto-oncogene products: members of
RT a new superfamily of enhancer binding proteins.";
RL Cold Spring Harb. Symp. Quant. Biol. 51:759-772(1988).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; M26747; AAA35677.1; -.
DR PFW; PF00104; hormone_rec; 1.
DR PFW; PF00105; zf-C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger.
SQ SEQUENCE 456 AA; 52176 MW; 3346388F CRC32;

```

Query Match 92.3%; Score 36; DB 4; Length 456;  
Best Local Similarity 83.3%; Pred No. 4.9;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
I:|||||  
Db 159 CKFKKC 164

```

RESULT 2
O00402 PRELIMINARY; PRT: 579 AA.
ID O00402
AC O00402
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE EAR-1R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA KAMIZONO A., SHUICHIRO K., KOHEI U.;

```

RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; D16815; BAA20088.1; -.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 SQ SEQUENCE 579 AA; 64643 MW; A8DCBE1B CRC32;

Query Match 92.3%; Score 36; DB 4; Length 579;  
 Best Local Similarity 83.3%; Pred. No. 5.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 I:||||  
 Db 159 CRFKKC 164

RESULT 3  
 Q14995 PRELIMINARY; PRT; 579 AA.  
 ID Q14995  
 AC Q14995  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE ORPHAN NUCLEAR HORMONE RECEPTOR BD73 (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PERIPHERAL BLOOD LYMPHOCYTES;  
 RX MEDLINE; 95089782.  
 RA DUMAS B., HARDING H.P., CHOI H.S., LEHMANN K.A., CHUNG M., LAZAR M.A.,  
 RA MOORE D.D.;  
 RT "A new orphan member of the nuclear hormone receptor superfamily  
 RT closely related to Rev-Erb.";  
 RL Mol. Endocrinol. 8:996-1005(1994).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL; L31785; AAA65937.1; -.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 FT NON\_TER  
 SQ SEQUENCE 579 AA; 64668 MW; E2F0ECC8 CRC32;

Query Match 92.3%; Score 36; DB 4; Length 579;  
 Best Local Similarity 83.3%; Pred. No. 5.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 I:||||  
 Db 158 CRFKKC 163

RESULT 4  
 O95779 PRELIMINARY; PRT; 445 AA.  
 ID O95779  
 AC O95779  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE BC39498.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA LAMERDIN J.E., MCCREARY P.M., SKOWRONSKI E., VISWANATHAN V.,  
 RA BURKHART-SCHULTZ K., GORDON L., DIAS J., RAMIREZ M., STILLWAGEN S.,  
 RA PHAN H., VELASCO N., DO L., REGALA W., TERRY A., GARNES J.,  
 RA DANGANAN L., ERLER A., CHRISTENSEN M., GEORGESCU A., AVILA J., LIU S.,  
 RA ATTIX C., ANDREISE T., FRANKHEIM M., AMICO-KELLER G., COEFIELD J.,  
 RA DUARTE S., LUCAS S., BRUCE R., THOMAS P., QUAN G., KRONMILLER B.,  
 RA ARELLANO A., SANDERS C., OW D., NOLAN M., TRONG S., KOBAYASHI A.,  
 RA OLSEN A.S., CARRANO A.V.;  
 RT "Sequence analysis of a 2 Mb region in 19p12 between UBA52 and  
 RT D19S455.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006539; AAD14470.1; -.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 10.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 445 AA; 51970 MW; 6CCFF9ED CRC32;

Query Match 92.3%; Score 36; DB 4; Length 445;  
 Best Local Similarity 83.3%; Pred. No. 4.8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 I:||||  
 Db 428 CKYKCC 433

RESULT 5  
 O95780 PRELIMINARY; PRT; 498 AA.  
 ID O95780  
 AC O95780  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE BC39498.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LAMERDIN J.E., MCCREARY P.M., SKOWRONSKI E., VISWANATHAN V.,  
 RA BURKHART-SCHULTZ K., GORDON L., DIAS J., RAMIREZ M., STILLWAGEN S.,  
 RA PHAN H., VELASCO N., DO L., REGALA W., TERRY A., GARNES J.,  
 RA DANGANAN L., ERLER A., CHRISTENSEN M., GEORGESCU A., AVILA J., LIU S.,  
 RA ATTIX C., ANDREISE T., FRANKHEIM M., AMICO-KELLER G., COEFIELD J.,  
 RA DUARTE S., LUCAS S., BRUCE R., THOMAS P., QUAN G., KRONMILLER B.,  
 RA ARELLANO A., SANDERS C., OW D., NOLAN M., TRONG S., KOBAYASHI A.,  
 RA OLSEN A.S., CARRANO A.V.;  
 RT "Sequence analysis of a 2 Mb region in 19p12 between UBA52 and  
 RT D19S455.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006539; AAD14472.1; -.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 10.  
 KW Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 498 AA; 58361 MW; 67CB5ABD CRC32;

Query Match 92.3%; Score 36; DB 4; Length 498;  
 Best Local Similarity 83.3%; Pred. No. 5.3;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CKFKKC 6  
 I:||||  
 Db 481 CKYKCC 486

RESULT 6  
 O18395 PRELIMINARY; PRT; 865 AA.  
 ID O18395  
 AC O18395  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE NUCLEAR HORMONE RECEPTOR.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON S;  
 RA RUSSELL S.R.H., HEIMBECK G., GODDARD C.M., CARPENTER A.T.C.,  
 RA ASHBURNER M.;  
 RL Genetics 144:156-170(1996).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: X98881; CAA67384.1; -;  
 DR FLYBASE: FBgn0004865; Eip78C.  
 DR PFAM: PF00104; hormone\_rec; 1.  
 DR PFAM: PF00105; zf-C4; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 KW Zinc-finger.  
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 SQ SEQUENCE 865 AA; 95179 MW; 219DFD6A CRC32;

Query Match 92.3%; Score 36; DB 5; Length 865;  
 Best Local Similarity 83.3%; Pred. No. 8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|:|:|  
 Db 422 CRFKKC 427

RESULT 7  
 ID 061228 PRELIMINARY; PRT; 810 AA.  
 AC 061228;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE THR4 GENE.  
 GN THR4.  
 OS Tenebrio molitor (Yellow mealworm).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Coleoptera; Polyphaga; Cucujiformia; Tenebrionidae;  
 OC Tenebrio.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MOUILLET J.F., SEDANO N., NICOLAI M., BOUSQUET F., LAUDET V.,  
 RA DELACHAMBRE J.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: AJ005685; CAA06670.1; -;  
 DR PFAM: PF00104; hormone\_rec; 1.  
 DR PFAM: PF00105; zf-C4; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 SQ SEQUENCE 810 AA; 89535 MW; 4F07D0A2 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 810;  
 Best Local Similarity 83.3%; Pred. No. 7.6;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|:|:|  
 Db 455 CRFKKC 460

RESULT 8  
 ID Q17929 PRELIMINARY; PRT; 389 AA.  
 AC Q17929;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE SIMILAR TO STEROID/THYROID/RETINOIC NUCLEAR HORMONE RECEPTORS.  
 GN C12D5.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA MILLER N., STELLYES L., BRADSHAW H.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL EMBL: U55365; AAA98573.1; -;  
 DR PFAM: PF00104; hormone\_rec; 1.  
 DR PFAM: PF00105; zf-C4; 1.  
 SQ SEQUENCE 389 AA; 45314 MW; F74C5317 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 389;  
 Best Local Similarity 83.3%; Pred. No. 4.4;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 |:|:|:|  
 Db 79 CRFKKC 84

RESULT 9  
 ID Q17013 PRELIMINARY; PRT; 556 AA.  
 AC Q17013;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE R13D11.8 PROTEIN.  
 GN R13D11.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,

RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans."; Pred. No. 5.7; Length 556;  
 RN Nature 368:32-38(1994).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA MILLER N., BECK C., GIBSON A.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF022983; AAB69942.1; -;  
 DR PFAM; PF00105; zf-C4; 1.  
 SQ SEQUENCE 556 AA; 62982 MW; F8840960 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 556;  
 Best Local Similarity 83.3%; Pred. No. 5.7;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFKKC 6  
 Db 125 CRFKKC 130

RESULT 10  
 O16609 PRELIMINARY; PRT; 363 AA.  
 ID O16609;  
 AC O16609;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE C49D10.2 PROTEIN.  
 GN C49D10.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans."; Pred. No. 5.7; Length 556;  
 DR EMBL; AF016665; AAC71177.1; -;  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 SQ SEQUENCE 363 AA; 41841 MW; 1D00695D CRC32;

Query Match 92.3%; Score 36; DB 5; Length 556;  
 Best Local Similarity 83.3%; Pred. No. 5.7;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFKKC 6  
 Db 125 CRFKKC 130

RESULT 10  
 O16609 PRELIMINARY; PRT; 363 AA.  
 ID O16609;  
 AC O16609;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE C49D10.2 PROTEIN.  
 GN C49D10.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF016665; AAC71177.1; -;  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 SQ SEQUENCE 363 AA; 41841 MW; 1D00695D CRC32;

Query Match 92.3%; Score 36; DB 5; Length 363;  
 Best Local Similarity 83.3%; Pred. No. 4.1;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFKKC 6  
 Db 66 CRFKKC 71

RESULT 11  
 O45460 PRELIMINARY; PRT; 422 AA.  
 ID O45460;  
 AC O45460;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE NHR-54 PROTEIN.  
 GN NHR-54.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MCMURRAY A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 94150718.  
 RX WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans."; Pred. No. 4.1;  
 DR EMBL; Z81531; CAB04316.1; -;  
 DR PFAM; PF01362; DUF12; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 SQ SEQUENCE 422 AA; 49473 MW; 5F0C0A13 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 422;  
 Best Local Similarity 83.3%; Pred. No. 4.6;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CKFKKC 6  
 Db 75 CRFKKC 80

RESULT 12  
 O17928 PRELIMINARY; PRT; 338 AA.  
 ID O17928;  
 AC O17928;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE NHR-52 PROTEIN.  
 GN NHR-52.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LLOYD C.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 283233; CAB05761.1; -.  
 SQ SEQUENCE 338 AA; 39582 MW; 0F9FB43C CRC32;

Query Match 92.3%; Score 36; DB 5; Length 338;  
 Best Local Similarity 83.3%; Pred. No. 3.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 58 CRFKKC 63

RESULT 13  
 O46003 PRELIMINARY; PRT; 342 AA.  
 AC O46003;  
 DT 01-JUN-1998 (TRENBLrel. 06, Created)  
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE ZK218.5 PROTEIN.  
 GN ZK218.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MCMURRAY A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 282085; CAB04989.1; -.  
 DR PFAM: PF00104; hormone\_rec; 1.  
 DR PFAM: PF00105; zf-C4; 1.  
 SQ SEQUENCE 342 AA; 39316 MW; DFC824E7 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 342;  
 Best Local Similarity 83.3%; Pred. No. 3.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 78 CRFKKC 83

RESULT 14  
 O45315 PRELIMINARY; PRT; 452 AA.  
 AC O45315;  
 DT 01-JUN-1998 (TRENBLrel. 06, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE C54E10.5 PROTEIN.  
 GN C54E10.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MATTHEWS L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: 281487; CAB03998.1; -.  
 DR PFAM: PF01362; DUF12; 1.  
 DR PFAM: PF00105; zf-C4; 1.  
 SQ SEQUENCE 452 AA; 53439 MW; F4F6AD68 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 452;  
 Best Local Similarity 83.3%; Pred. No. 4.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 74 CRFKKC 79

RESULT 15  
 O56836 PRELIMINARY; PRT; 1464 AA.  
 AC O56836;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE EG-133E12.2 PROTEIN.  
 GN EG-133E12.2.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;  
 OC Drosophilidae; Drosophila.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA FERRAZ C., VIDAL S., BRUN C., BUCHETON A., DEMAILLE J.G.;  
 RT "Sequencing the distal X chromosome of Drosophila melanogaster";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA BENOS P.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: AL035245; CAA22836.1; -.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 KW Receptor; transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 SQ SEQUENCE 1464 AA; 158008 MW; 2A110783 CRC32;

Query Match 92.3%; Score 36; DB 5; Length 1464;  
 Best Local Similarity 83.3%; Pred. No. 12;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKKC 6  
 DB 921 CRFKKC 926

Search completed: September 7, 1999, 20:34:46  
Job time: 19742 sec

---



GenCore version 4.5

Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:16 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-15

Perfect score: 61

Sequence: 1 KFKCKFKFKC 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	10	1 R71786	Peptide neutralisi
2	61	100.0	10	1 W21603	Antibiotic potenti
3	43	70.5	10	1 R71787	Peptide neutralisi
4	43	70.5	10	1 W21604	Antibiotic potenti
5	41	67.2	10	1 R71776	Peptide neutralisi
6	41	67.2	10	1 W21593	Antibiotic potenti
7	39.5	64.8	10	1 R71789	Peptide neutralisi
8	39.5	64.8	10	1 W21606	Peptide neutralisi
9	37	60.7	10	1 R33526	Peptide for treati
10	37	60.7	10	1 R33531	Peptide for treati
11	37	60.7	11	1 R33288	Endotoxin lipid A
12	37	60.7	11	1 R33292	Endotoxin lipid A
13	37	60.7	10	1 W21619	Antibiotic potenti
14	37	60.7	11	1 W21623	Antibiotic potenti
15	37	60.7	10	1 W62434	Human neutrophil g
16	37	60.7	11	1 W62435	Human neutrophil g
17	36	59.0	10	1 W21631	Antibiotic potenti
18	35	57.4	652	1 W48879	Amino acid sequenc
19	34	55.7	263	1 P92003	Deduced sequence o
20	33	54.1	1221	1 R52699	Sequence translate
21	33	54.1	285	1 R64810	ScFv anti-lysozyme
22	33	54.1	4655	1 R97208	Human calcium sens
23	33	54.1	4855	1 R97209	Human placental ca
24	33	54.1	4855	1 R97210	Human kidney calci
25	33	54.1	4655	1 R97211	Human parathyroid
26	33	54.1	107	1 W19261	Lactobacillus amyl
27	33	54.1	107	1 W19264	Lactobacillus amyl
28	33	54.1	330	1 W34478	Zfy protein. zfy a
29	33	54.1	4855	1 W43311	Human calcium sens
30	33	54.1	4855	1 W43312	Human placental ca
31	33	54.1	4855	1 W43313	Human kidney calci
32	33	54.1	4655	1 R43314	Human parathyroid
33	32	52.5	68	1 R08246	GTG32 chimeric glu
34	32	52.5	68	1 R08248	GTG33 chimeric glu
35	32	52.5	58	1 R08249	GTG28 chimeric glu
36	32	52.5	456	1 P80921	Sequence encoded b
37	32	52.5	3080	1 R35081	ZYM polypeptide
38	32	52.5	514	1 R47621	Rat thyroid hormon
39	32	52.5	10	1 R71788	Peptide neutralisi
40	32	52.5	88	1 R90991	Mouse Wat-8 polype
41	32	52.5	10	1 W21605	Antibiotic potenti
42	32	52.5	941	1 W55469	H. pylori ORF 05ee
43	32	52.5	548	1 W56014	Recombinant botuli

## ALIGNMENTS

## RESULT 1

R71786  
ID R71786 standard; peptide; 10 AA.  
AC R71786;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 4..10  
PN WO9503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
repeating units of a basic aminoacid or basic and hydrophobic  
amino acids  
PS Claim 17; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides which  
include units of formula: (A)n, where A is the cationic amino acid Lys  
or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
(ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
Trp, and p is 2 or greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
activity. Hence they can be used therapeutically to treat septic shock  
and also in vitro to detoxify vaccines, drug solutions, injectable  
nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
peptides. 10 AA;  
SQ Sequence 10 AA;

Query Match 100.0%; Score 61; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0005;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10

|||||||

DB 1 KFKCKFKFKC 10

## RESULT 2

W21603  
ID W21603 standard; peptide; 10 AA.  
AC W21603;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #15.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 4..10  
PN WO9638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M.  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
amino acid sequence - reduces dose of antibiotic required  
PS Claim 19; Page 25; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 61; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0005;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
 Db 1 KFKCKFKFKC 10  
 |||||

RESULT 3  
 R71787  
 ID R71787 standard; peptide; 10 AA.  
 AC R71787;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 DT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 18; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 70.5%; Score 43; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.27;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
 Db 1 KFKCKFKFKC 10  
 |||||

RESULT 4  
 W21604  
 ID W21604 standard; peptide; 10 AA.  
 AC W21604;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #16.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.  
 FH Synthetic.  
 OS Key Location/Qualifiers  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M; Varra M;  
 DR WPI; 97-034095/03.  
 DT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 20; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 70.5%; Score 43; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.27;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
 Db 1 KFKCKFKFKC 10  
 |||||

RESULT 5  
 R71776  
 ID R71776 standard; peptide; 10 AA.  
 AC R71776;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 DT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 7; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater..  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 67.2%; Score 41; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.54;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKCKFKFK 9  
 DB 1 KFKCKFKFK 9

## RESULT 6

W21593  
 ID W21593 standard; peptide; 10 AA.  
 AC W21593;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #5.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 9; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 67.2%; Score 41; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.54;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKCKFKFK 9  
 DB 1 KFKCKFKFK 9

## RESULT 7

R71789  
 ID R71789 standard; peptide; 10 AA.  
 AC R71789;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 2..9  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-057830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic aminoacid or basic and hydrophobic  
 PT amino acids  
 PS Claim 20; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic

CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 64.8%; Score 39.5; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.92;  
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3 KCKFK-KFKC 10  
 DB 1 KCKFKKFKC 9

## RESULT 8

W21606  
 ID W21606 standard; peptide; 10 AA.  
 AC W21606;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #18.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 2..9  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 22; Page 26; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 64.8%; Score 39.5; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.92;  
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3 KCKFK-KFKC 10  
 DB 1 KCKFKKFKC 9

## RESULT 9

R33526  
 ID R33526 standard; peptide; 10 AA.  
 AC R33526;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 OS HIV related infections; polymyxin B.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN ZA9200943-A.  
 PD 25-NOV-1992.

PF 10-FEB-1992: 000943.  
 PR 11-FEB-1991: US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 4: Page 32: 39pp; English.  
 CC This is a specific example of a generic peptide of formula  
 CC R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where n = 1-100  
 CC and each R is H, an amino acid residue or a fatty acid residue.  
 CC The peptide is useful for treating or preventing septic shock.  
 CC mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 10 AA;

Query Match 60.7%; Score 37; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2.2;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKEKFKC 10  
 | | | | | |  
 DB 1 KTKCKFLKCC 10

RESULT 10  
 R33531  
 ID R33531 standard; peptide; 11 AA.  
 AC R33531; 1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992: 000943.  
 PR 11-FEB-1991: US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 8: Page 32: 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 60.7%; Score 37; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 2.4;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKEKFKC 10  
 | | | | | |  
 DB 2 KTKCKFLKCC 11

RESULT 11  
 R39288  
 ID R39288 standard; peptide; 10 AA.  
 AC R39288;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymixin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992: E01060.  
 PR 16-JAN-1992: US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 4: Page 32: 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 10 AA;

Query Match 60.7%; Score 37; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2.2;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKEKFKC 10  
 | | | | | |  
 DB 1 KTKCKFLKCC 10

RESULT 12  
 R39292  
 ID R39292 standard; peptide; 11 AA.  
 AC R39292;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymixin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;

quantification.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN W09314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 8; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins; and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 60.7%; Score 37; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 2.4;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
 | | | | |  
 Db 2 KTKCKFLKCC 11

## RESULT 13

W21619  
 ID W21619 standard; peptide; 10 AA.  
 AC W21619;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #31.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 4..10  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 35; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 60.7%; Score 37; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 2.2;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 KFKCKFKFKC 10  
 | | | | |  
 Db 1 KTKCKFLKCC 10  
 RESULT 14  
 W21623  
 ID W21623 standard; peptide; 11 AA.  
 AC W21623;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #35.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 39; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 60.7%; Score 37; DB 1; Length 11;  
 Best Local Similarity 70.0%; Pred. No. 2.4;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
 | | | | |  
 Db 2 KTKCKFLKCC 11

## RESULT 15

W62434  
 ID W62434 standard; peptide; 10 AA.  
 AC W62434;  
 DT 01-OCT-1998 (first entry)  
 DE Human neutrophil granule bactericidal derived peptide SEQ ID NO:49.  
 KW Bactericidal; Pseudomonas sp.; endotoxin shock; bacterial infection;  
 KW bactericidal permeability increasing factor; B/PI; neutralising;  
 KW human neutrophil granule bactericidal protein.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN US5786324-A.  
 PD 28-JUL-1998.  
 PF 24-MAR-1994; 218026.  
 PR 24-MAR-1994; US-218026.  
 PA (MINO ) UNIV MINNESOTA.  
 PI Gray B, Haseman JR, Mayo K;  
 DR WPI: 98-436578/37.  
 PT Bactericidal and endotoxin-neutralising peptides - used in treating  
 PT e.g. Pseudomonas species infection and in protectively coating  
 PT prosthetic devices  
 PS Example 1; Column 18; 46pp; English.  
 CC The present invention describes bactericidal peptides (BP); (a) for  
 CC Pseudomonas species; (b) with endotoxin neutralising activity; (c)

CC with both endotoxin neutralising activity and bactericidal activity for  
CC Pseudomonas aeruginosa; (d) a hybrid peptide with endotoxin neutralising  
CC activity and bactericidal activity, the peptide having at least 1 amino  
CC acid sequence of a peptide with endotoxin neutralising activity combined  
CC with at least 1 different amino acid sequence of a peptide that has  
CC bactericidal activity; and (e) a prosthetic device which has a  
CC sufficient amount of BP attached to the surface to inhibit bacterial  
CC growth. The peptides of the invention are used in treating bacterial  
CC infection such as Pseudomonas strains e.g. for P. aeruginosa at 10<sup>-7</sup> to  
CC 10<sup>-9</sup> M, and Escherichia coli. The peptides are also used to treat  
CC endotoxin shock. The present sequence represents a peptide derived  
CC from human neutrophil granule bactericidal protein from an example of  
CC the present invention.  
SQ Sequence 10 AA;

Query Match 60.7%; Score 37; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 2.2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KFKCKFRFKC 10  
| | | | | |  
Db 1 KTKCKFLKKC 10

Search completed: September 7, 1999, 20:37:17  
Job time: 18466 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:37 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-15

Perfect score: 61

Sequence: 1 KFKCKFKFKC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	10	1	US-08-097-830E-15
2	61	100.0	10	2	US-08-456-112B-15
3	43	70.5	10	1	US-08-097-830E-16
4	43	70.5	10	2	US-08-456-112B-16
5	41	67.2	10	1	US-08-097-830E-5
6	41	67.2	10	2	US-08-456-112B-5
7	39.5	64.8	10	1	US-08-097-830E-18
8	39.5	64.8	10	2	US-08-456-112B-18
9	37	60.7	10	1	US-08-049-871-2
10	37	60.7	11	1	US-08-049-871-6
11	37	60.7	10	1	US-07-819-893-2
12	37	60.7	11	1	US-07-819-893-6
13	37	60.7	10	1	US-08-280-397-2
14	37	60.7	11	1	US-08-280-397-6
15	37	60.7	10	2	US-08-218-026-49
16	37	60.7	11	2	US-08-218-026-50
17	37	60.7	10	2	US-08-653-632-49
18	37	60.7	11	2	US-08-653-632-50
19	37	60.7	11	2	US-08-456-112B-13
20	37	60.7	10	2	US-08-456-112B-31
21	37	60.7	11	2	US-08-456-112B-35
22	36	59.0	10	2	US-08-456-112B-43
23	34	55.7	263	1	US-07-906-983-2
24	32	52.5	10	1	US-08-097-830E-17
25	32	52.5	10	2	US-08-456-112B-17
26	32	52.5	3224	2	US-08-703-660-34
27	31	50.8	1003	1	US-07-991-867B-6
28	31	50.8	249	1	US-08-155-171B-32
29	31	50.8	64	1	US-08-344-695-11
30	31	50.8	10	1	US-08-097-830E-12
31	31	50.8	1003	1	US-08-107-755A-6
32	31	50.8	10	2	US-08-456-112B-12
33	31	50.8	142	2	US-08-448-438-1
34	31	50.8	142	2	US-08-448-438-2
35	31	50.8	142	2	US-08-448-438-3
36	31	50.8	142	2	US-08-448-438-4
37	31	50.8	160	2	US-08-448-438-5
38	31	50.8	160	2	US-08-448-438-6
39	31	50.8	160	2	US-08-448-438-7

Sequence 8, Appli  
Sequence 1, Appli  
Sequence 4, Appli  
Sequence 5, Appli  
Sequence 1, Appli  
Patent No. 5206163

#### ALIGNMENTS

##### RESULT 1

US-08-097-830E-15  
; Sequence 15, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and  
; OTHER INFORMATION: Cys  
US-08-097-830E-15

Query Match 100.0%; Score 61; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00069;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFKCKFKFKC 10

DB 1 KFKCKFKFKC 10

##### RESULT 2

US-08-456-112B-15  
; Sequence 15, Application US/08456112B  
; Patent No. 5634430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-15

Query Match 100.0%; Score 61; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00069;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
| | | | | | | |  
Db 1 KFKCKFKFKC 10

RESULT 3  
US-08-097-830E-16  
Sequence 16, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE: sulfide bond between Cys and Cys  
US-08-097-830E-16

Query Match 70.5%; Score 43; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.27;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
| | | | | | | |  
Db 1 KFKCKFKFKC 10

RESULT 4  
US-08-456-112B-16  
Sequence 16, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-16

Query Match 70.5%; Score 43; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.27;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
| | | | | | | |



```
Db      1 KLKCKLKLC 10

RESULT 5
US-08-097-830E-5
; Sequence 5, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-097-830E-5

Query Match      67.2%; Score 41; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KFKCKFKFK 9
Db      1 KFKCKFKFK 9

RESULT 7
US-08-097-830E-18
; Sequence 18, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular

Query Match      67.2%; Score 41; DB 2; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KFKCKFKFK 9
Db      1 KFKCKFKFK 9

RESULT 6
US-08-456-112B-5
; Sequence 5, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486

Query Match      67.2%; Score 41; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 KFKCKFKFK 9
Db      1 KFKCKFKFK 9

RESULT 6
US-08-456-112B-5
; Sequence 5, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
```

FEATURE:  
OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-18

Query Match 64.8%; Score 39.5; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.85;  
Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3 KCKF-KFKC 10  
|||||  
Db 1 KCKFKFKC 9

## RESULT 8

US-08-456-112B-18  
Sequence 18, Application US/08456112B  
Patent No. 5834430

GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995

## CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-456-112B-18

Query Match 64.8%; Score 39.5; DB 2; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.85;  
Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3 KCKF-KFKC 10  
|||||  
Db 1 KCKFKFKC 9

## RESULT 9

US-08-049-871-2

Sequence 2, Application US/08049871

Patent No. 5358933

GENERAL INFORMATION:

APPLICANT: Porro, Massimo

TITLE OF INVENTION: Synthetic Peptides for Detoxification

TITLE OF INVENTION: Of Bacterial Endotoxins and for the

TITLE OF INVENTION: Prevention and Treatment of Septic  
Shock  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/049,871  
FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/658,744

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-049-871-2

Query Match 60.7%; Score 37; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10  
| | | | |  
Db 1 KTKCKFKKCK 10

## RESULT 10

US-08-049-871-6

Sequence 6, Application US/08049871

Patent No. 5358933

GENERAL INFORMATION:

APPLICANT: Porro, Massimo

TITLE OF INVENTION: Synthetic Peptides for Detoxification

TITLE OF INVENTION: Of Bacterial Endotoxins and for the

TITLE OF INVENTION: Prevention and Treatment of Septic

TITLE OF INVENTION: Shock

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson, Costigan & Hoare

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,871

FILING DATE:

CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/658,744  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-049-871-6

Query Match 60.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 2.1;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKKC 10  
| | | | | | |  
Db 2 KTKCKFLKC 11

RESULT 11  
US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porrio, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-2

Query Match 60.7%; Score 37; DB 1; Length 10;

Best Local Similarity 70.0%; Pred. No. 2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKKC 10  
| | | | | | |  
Db 1 KTKCKFLKC 10

RESULT 12  
US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porrio, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
; US-07-819-893-6

Query Match 60.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 2.1;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKKC 10  
| | | | | | |  
Db 2 KTKCKFLKC 11

RESULT 13  
US-08-280-397-2  
; Sequence 2, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porrio, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;;  
;; US-08-280-397-2

Query Match 50.7%; Score 37; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
1 ||||| 11  
DB 1 KTKCKFLKCC 10

RESULT 14  
US-08-280-397-6  
;; Sequence 6, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;;  
;; US-08-280-397-6

Query Match 60.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 70.0%; Pred. No. 2.1;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
1 ||||| 11  
DB 2 KTKCKFLKCC 11

RESULT 15  
US-08-218-026-49  
;; Sequence 49, Application US/08218026  
;; Patent No. 5786324  
;; GENERAL INFORMATION:  
;; APPLICANT: Gray, Beulah  
;; APPLICANT: Haseman, Judith R.  
;; APPLICANT: Mayo, Kevin  
;; TITLE OF INVENTION: Synthetic Peptides with Bactericidal  
;; TITLE OF INVENTION: Activity and Endotoxin Neutralizing Activity for Gram  
;; TITLE OF INVENTION: Negative Bacteria and Methods for Their Use  
;; NUMBER OF SEQUENCES: 60  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merchant & Gould  
;; STREET: 3100 No. 5786324west Center  
;; CITY: Minneapolis  
;; STATE: MN  
;; COUNTRY: USA  
;; ZIP: 55402  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PG-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/218,026  
;; FILING DATE: 24-MAR-1994  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kowalchuk, Katherine M.  
;; REGISTRATION NUMBER: 36,848  
;; REFERENCE/DOCKET NUMBER: 600.286US01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 612-332-5300  
;; TELEFAX: 612-332-9081  
;; INFORMATION FOR SEQ ID NO: 49:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;;  
;; US-08-218-026-49

Query Match 60.7%; Score 37; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 2;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10

Db 1 1111 11  
1 KTKCKLKC 10

Search completed: September 7, 1999, 22:38:37  
Job time: 7921 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:28 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-15

Perfect score: 61  
Sequence: 1 KFKCKFKFC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	68.9	137	2 S68429	myotoxin precursor
2	41	67.2	170	2 C71199	hypothetical prote
3	39.5	64.8	1060	2 S33641	homeothetic protein z
4	36	59.0	121	1 PSSNAM	phospholipase A2 h
5	36	59.0	42	1 MXRSMV	myotoxin a 6 - pra
6	36	59.0	43	1 CXRSCH	toxic peptide C -
7	36	59.0	291	2 S10329	aryl sulfotransfer
8	36	59.0	43	2 B29089	myotoxin II - midg
9	36	59.0	43	2 A29089	myotoxin I - midg
10	36	59.0	45	2 S12909	myotoxin - western
11	36	59.0	65	2 JC5324	myotoxin a precursor
12	36	59.0	398	2 D69953	conserved hypothet
13	36	59.0	363	2 C69255	first mannosiyl tra
14	35	57.4	1068	2 S64015	pleiotropic drug r
15	35	57.4	345	2 S23862	hypothetical prote
16	35	57.4	2616	2 A57096	nudel protein prec
17	34	55.7	263	1 WMVZSP	apolipoprotein H h
18	34	55.7	373	1 PDBYB	REPI protein - yea
19	34	55.7	380	2 S11187	REPI protein - yea
20	34	55.7	382	2 S40417	REPI protein - yea
21	34	55.7	1361	2 S50943	hypothetical prote
22	34	55.7	1817	2 H71611	probable secreted
23	34	55.7	2496	2 A71616	secreted protein P
24	34	55.7	435	2 S40993	hypothetical prote
25	34	55.7	568	2 A34891	Ig heavy chain pre
26	33	54.1	42	1 CXRSMT	crotonamine - tropic
27	33	54.1	51	2 D35947	crotonamine 4 precu
28	33	54.1	45	2 A37909	myotoxin - eastern
29	33	54.1	40	2 C39560	myotoxin a 5 - pra
30	33	54.1	934	2 A34372	complement C6 prec
31	33	54.1	876	2 A53998	late expression fa
32	33	54.1	143	2 G71655	probable sigma (54
33	33	54.1	416	2 H71433	hypothetical prote
34	33	54.1	757	2 S27762	Sipi protein - bar
35	33	54.1	940	2 A40985	projectin - fruit
36	33	54.1	1650	2 S53457	dominant autoantig
37	32	52.5	856	1 G64596	DNA ligase (NAD+)
38	32	52.5	1291	1 A48940	bontoxilysin (EC 3
39	32	52.5	456	1 TVHUAR	thyroid hormone re

40 32 52.5 369 1 TVCHTB thyroid hormone re  
41 32 52.5 138 2 S10992 ammodytin L precu  
42 32 52.5 122 2 S19570 ammodytin L - sand  
43 32 52.5 637 2 S42251 probable nucleosid  
44 32 52.5 656 2 A71916 DNA ligase - Helic  
45 32 52.5 1291 2 I40631 non-proteolytic bo

#### ALIGNMENTS

RESULT 1  
S68429  
myotoxin precursor - southern copperhead  
C:Species: Agkistrodon contortrix contortrix (southern copperhead)  
C>Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 13-Sep-1998  
C:Accession: S68429; S74296  
R:de Araujo, H.S.S.; White, S.P.; Ownby, C.L.  
Arch. Biochem. Biophys. 326, 21-30, 1996  
A:Title: cDNA cloning and sequence analysis of a lysine-49 phospholipase A(2) myotoxi  
A:Reference number: S68429  
A:Accession: S68429  
A:Molecule type: mRNA  
A:Residues: 1-137 <DER>  
A:Cross-references: EMBL:U21335; NID:g809484; PID:g809485  
A:Accession: S74296  
A:Molecule type: protein  
A:Residues: 17-36 <DER>  
C:Superfamily: phospholipase A2  
F:i-16/Domain: signal sequence #status predicted <SIG>  
F:i-137/Product: myotoxin #status experimental <MAT>

Query Match 68.9% Score 42; DB 2; Length 137;  
Best Local Similarity 70.0%; Pred. No. 2.1;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
||| |||||  
DB 122 KYKAYFKFC 131

RESULT 2  
C71199  
hypothetical protein PH1864 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998  
C:Accession: C71199  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Ogu  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil  
A:Reference number: A71000; MUID:98344137  
A:Accession: C71199  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-170 <KAW>  
A:Cross-references: GB:AP000007; NID:g3236134; PID:d1031929; PID:g3258303  
A:Experimental source: strain OY3  
A>Note: this accession replaces an interim accession for a sequence replaced by GenBa  
C:Genetics:  
A:Gene: PH1864

Query Match 67.2% Score 41; DB 2; Length 170;  
Best Local Similarity 70.0%; Pred. No. 3.6;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
||||| |  
DB 16 KFKCKFCFLDC 25

```

RESULT 3
S33641
homeotic protein zfh-1 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1998
C:Accession: S33641; S27816
R:Fortini, M.E.; Lai, Z.; Rubin, G.M.
Mech. Dev. 34, 113-122, 1991
A:Title: The Drosophila zfh-1 and zfh-2 genes encode novel proteins containing both zinc
A:Reference number: S33641
A:Accession: S33641
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1060 <FOR>
A:Cross-references: EMBL:M63449; NID:gl158820; PID:gl158821
C:Genetics:
A:Gene: zfh-1
A:Cross-references: FlyBase:FBgn0004606
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation; zinc finger
F:706-762/Domain: homeobox homology <HOX>

Query Match 64.8%; Score 39.5; DB 2; Length 1060;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 KFKCK-----FKFK 9
||||| |||||
DB 360 KFKCKECKAKFKK 373

RESULT 4
PSSNAM
Phospholipase A2 homolog, inactive - eastern cottonmouth
C:Species: Agkistrodon piscivorus piscivorus (eastern cottonmouth)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 14-Feb-1997
C:Accession: A00767; A60719
R:Maraganore, J.M.; Heinrikson, R.L.
J. Biol. Chem. 261, 4797-4804, 1986
A:Title: The lysine-49 phospholipase A-2 from the venom of Agkistrodon piscivorus piscivorus
A:Reference number: A00767; MUID:86168190
A:Accession: A00767
A:Molecule type: protein
A:Residues: 1-121 <MAR>
R:van den Bergh, C.J.; Slotboom, A.J.; Verheij, H.M.; de Haas, G.H.
J. Cell. Biochem. 39, 379-390, 1989
A:Title: The role of Asp-49 and other conserved amino acids in phospholipases A2 and the
A:Reference number: A60719; MUID:89255682
A:Accession: A60719
A:Molecule type: protein
A:Residues: 1-8 <VAN>
A:Note: The authors confirmed the identity of the purified protein by composition and pe
R:Holland, D.R.; Clancy, L.L.; Muchmore, S.W.; Rydel, T.J.; Einspahr, H.M.; Finzel, B.C.
submitted to the Brookhaven Protein Data Bank, October 1991
A:Reference number: A51920; PDB:1PPA
R:Holland, D.R.; Clancy, L.L.; Muchmore, S.W.; Ryde, T.J.; Einspahr, H.M.; Finzel, B.C.;
J. Biol. Chem. 265, 17649-17656, 1990
A:Title: The crystal structure of a lysine 49 phospholipase A-2 from the venom of the co
A:Reference number: A39221; MUID:91009222
A:Contents: annotation; X-ray crystallography, 2.0 angstroms
C:Comment: This homolog of phospholipase A2 lacks enzymatic activity because a Lys-48 re
C:Superfamily: phospholipase A2
C:Keywords: venom
F:26-115,28-44,43-95,49-121,50-88,57-81,75-86/disulfide bonds: #status experimental

Query Match 59.0%; Score 36; DB 1; Length 121;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKK 10

```

```

DB 106 KYKAYFKLKC 115
|:| | | | | |
RESULT 5
MARSMW
myotoxin a 6 - prairie rattlesnake
C:Species: Crotalus viridis viridis (prairie rattlesnake)
C:Date: 30-Apr-1979 #sequence_revision 30-Apr-1979 #text_change 02-Jul-1998
C:Accession: A01736; S27115; A39560; JC5817
R:Fox, J.W.; Elzinga, M.; Tu, A.T.
Biochemistry 18, 678-684, 1979
A:Title: Amino acid sequence and disulfide bond assignment of myotoxin a isolated fro
A:Reference number: A01736; MUID:79124714
A:Accession: A01736
A:Molecule type: protein
A:Residues: 1-42 <FOX>
R:Baker, B.; Utalsincharoen, P.; Tu, A.T.
Arch. Biochem. Biophys. 298, 325-331, 1992
A:Title: Structure-function relationship of myotoxin a using peptide fragments.
A:Reference number: S27115; MUID:93037463
A:Accession: S27115
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-42 <BAK>
R:Aird, S.D.; Kruggel, W.G.; Kaiser, I.I.
Toxicol. 29, 265-268, 1991
A:Title: Multiple myotoxin sequences from the venom of a single prairie rattlesnake (
A:Reference number: A39560; MUID:91263105
A:Accession: A39560
A:Molecule type: protein
A:Residues: 1-42 <AIR>
R:Nedelkov, D.; O'Keefe, M.P.; Chapman, T.L.; Bieber, A.L.
Biochem. Biophys. Res. Commun. 241, 525-529, 1997
A:Title: The role of Pro20 in the isomerization of myotoxin a from Crotalus viridis v
A:Reference number: JC5817
A:Accession: JC5817
A:Molecule type: protein
A:Residues: 1-42 <NED>
C:Comment: This protein inhibits Ca2+ uptake into the sarcoplasmic reticulum by bindi
C:Superfamily: crotoxin
C:Keywords: myotoxin
F:4-36,11-30,18-37/disulfide bonds: #status experimental

Query Match 59.0%; Score 36; DB 1; Length 42;
Best Local Similarity 50.0%; Pred. No. 7.2;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFKK 10
| | | | | | |
DB 27 KMDCRWKKC 36

RESULT 6
CXRSCH
toxic peptide C - southern Pacific rattlesnake
C:Species: Crotalus viridis helleri (southern Pacific rattlesnake)
C:Date: 31-May-1979 #sequence_revision 08-Oct-1981 #text_change 23-Aug-1996
C:Accession: A01737
R:Maeda, N.; Tamiya, N.; Pattabhiraman, T.R.; Russell, F.E.
Toxicol. 16, 431-441, 1978
A:Title: Some chemical properties of the venom of the rattlesnake, Crotalus viridis h
A:Reference number: A01737; MUID:79015339
A:Accession: A01737
A:Molecule type: protein
A:Residues: 1-43 <MAE>
C:Superfamily: crotoxin
C:Keywords: myotoxin; venom
F:4-36,11-30,18-37/disulfide bonds: #status predicted

Query Match 59.0%; Score 36; DB 1; Length 43;

```



```
Best Local Similarity 50.0%; Pred. No. 7.3;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10
   | | | | |
Db 27 KMDCRWKWK 36

RESULT 7
aryl sulfotransferase (EC 2.8.2.1) IV - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Aug-1998
C:Accession: S10329; A44932
R:Ozawa, S.; Nagata, K.; Gong, D.; Yamazoe, Y.; Kato, R.
Nucleic Acids Res. 18, 4001, 1990
A:Title: Nucleotide sequence of a full-length cDNA (PST-1) for aryl sulfotransferase fr
A:Reference number: S10329; MUID:90326537
A:Accession: S10329
A:Molecule type: mRNA
A:Residues: 1-291 <OZA>
A:Cross-references: EMBL:X52883; NID:g55760; PID:g55761
R:Yerokun, T.; Etheredge, J.L.; Norton, T.R.; Carter, H.A.; Chung, K.H.; Birckbichler, F
Cancer Res. 52, 4779-4786, 1992
A:Title: Characterization of a complementary DNA for rat liver aryl sulfotransferase IV
uced hepatocarcinogenesis.
A:Reference number: A44932; MUID:92379794
A:Accession: A44932
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 7-291 <YEA>
A:Cross-references: GB:X68640; GB:S42994; NID:g55764; PID:g55765
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:112171, NCBIP:112178)
C:Superfamily: alcohol sulfotransferase
C:Keywords: sulfotransferase

Query Match 59.0%; Score 36; DB 2; Length 291;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CKFKFKC 10
   | | | | |
Db 283 CDFKFKC 289

RESULT 8
myotoxin II - midget faded rattlesnake
C:Species: Crotalus viridis concolor (midget faded rattlesnake)
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 11-Nov-1994
C:Accession: B29089
R:Bieber, A.L.; McParland, R.H.; Becker, R.R.
Toxicon 25, 677-680, 1987
A:Title: Amino acid sequences of myotoxins from Crotalus viridis concolor venom.
A:Reference number: A94323; MUID:87320413
A:Accession: B29089
A:Molecule type: protein
A:Residues: 1-43 <BIE>
A:Note: 15-Thr, 16-Val, 19-Leu, and 33-Arg were also found
C:Superfamily: crotamine
C:Keywords: myotoxin

Query Match 59.0%; Score 36; DB 2; Length 43;
Best Local Similarity 50.0%; Pred. No. 7.3;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10
   | | | | |
Db 27 KMDCRWKWK 36

RESULT 9
myotoxin I - midget faded rattlesnake
C:Species: Crotalus viridis concolor (midget faded rattlesnake)
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 11-Nov-1994
C:Accession: A29089
R:Bieber, A.L.; McParland, R.H.; Becker, R.R.
Toxicon 25, 677-680, 1987
A:Title: Amino acid sequences of myotoxins from Crotalus viridis concolor venom.
A:Reference number: A94323; MUID:87320413
A:Accession: A29089
A:Molecule type: protein
A:Residues: 1-43 <BIE>
C:Superfamily: crotamine
C:Keywords: myotoxin

Query Match 59.0%; Score 36; DB 2; Length 43;
Best Local Similarity 50.0%; Pred. No. 7.3;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10
   | | | | |
Db 27 KMDCRWKWK 36

RESULT 10
myotoxin - western rattlesnake
C:Species: Crotalus viridis (western rattlesnake)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S12909
R:Griffin, P.R.; Aird, S.D.
FEBS Lett. 274, 43-47, 1990
A:Title: A new small myotoxin from the venom of the prairie rattlesnake (Crotalus vir
A:Reference number: S12909; MUID:91071443
A:Accession: S12909
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-45 <GRI>
A:Note: 8-Gly was also found
C:Superfamily: crotamine

Query Match 59.0%; Score 36; DB 2; Length 45;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10
   | | | | |
Db 27 KMDCRWKWK 36

RESULT 11
myotoxin a precursor - prairie rattlesnake
C:Species: Crotalus viridis viridis (prairie rattlesnake)
C:Date: 15-May-1997 #sequence_revision 18-Jul-1997 #text_change 02-Jul-1998
C:Accession: JC5324
R:Norris, J.W.; Fry, R.M.; Tu, A.T.
Biochem. Biophys. Res. Commun. 230, 607-610, 1997
A:Title: The nucleotide sequence of the translated and untranslated regions of a cDNA
A:Reference number: JC5324; MUID:97167753
A:Accession: JC5324
A:Molecule type: mRNA
A:Residues: 1-65 <NOR>
A:Experimental source: venom
A:Note: the authors translated the codon CAG for residue 27 as His and CTT for residu
C:Comment: this protein inhibits Caz+ uptake into the sarcoplasmic reticulum by bindi
C:Superfamily: crotamine
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-64/Product: myotoxin a #status predicted <MAT>
```

Query Match 59.0%; Score 36; DB 2; Length 65;  
 Best Local Similarity 50.0%; Pred. No. 10;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
 | | | | |  
 DB 49 KMDCRWKRC 58

RESULT 12  
 D69953  
 conserved hypothetical protein yqfD - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
 C:Accession: D69953  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: D69953  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-398 <KUN>  
 A:Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PID:g1183765; PID:g2634981  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yqfD

Query Match 59.0%; Score 36; DB 2; Length 398;  
 Best Local Similarity 75.0%; Pred. No. 45;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFKCKFKF 8  
 | | | | |  
 DB 64 KFKCKARF 71

RESULT 13  
 C69255  
 first mannosyl transferase (wba2-1) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
 C:Accession: C69255  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: C69255  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-363 <KLE>  
 A:Cross-references: GB:AE001104; GB:AE0000782; NID:g2689427; PID:g2650610; TIGR:AF0043

Query Match 59.0%; Score 36; DB 2; Length 363;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KFKCKFC 10  
 | | | | |  
 DB 177 KFKCKFC 182

RESULT 14  
 S64015  
 pleiotropic drug resistance control protein PDR1 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein G372; protein YGL013c  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 06-Feb-1998  
 C:Accession: S64015; S31560; A28466  
 R:Hebling, U.; Hofmann, B.; Delli, H.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64003  
 A:Accession: S64015  
 A:Molecule type: DNA  
 A:Residues: 1-1068 <HEB>  
 A:Cross-references: EMBL:Z72535; NID:gl322470; PID:e243845; PID:gl322471; MIPS:YGL013  
 R:Experimental source: strain S288C  
 R:Chen, W.; Balzi, E.; Capieaux, E.; Choder, M.; Goffeau, A.  
 Yeast 7, 287-299, 1991  
 A:Title: The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1 and ATE1  
 ntrolling pleiotropic drug resistance.  
 A:Reference number: S15040; MUID:91353083  
 A:Accession: S31560  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-410 'K', 412-529, 'R', 531-819, 'A', 821-920, 'I', 922-980, 'S', 982-1009, 1015-1  
 R:Balzi, E.; Chen, W.; Ulaszewski, S.; Capieaux, E.; Goffeau, A.  
 J. Biol. Chem. 262, 16871-16879, 1987  
 A:Title: The multidrug resistance gene PDR1 from Saccharomyces cerevisiae.  
 A:Reference number: A28466; MUID:88059084  
 A:Accession: A28466  
 A:Molecule type: DNA  
 A:Residues: 1-410 'K', 412-819, 'A', 821-920, 'I', 922-980, 'S', 982-1009, 1015-1068 <BAL>  
 A:Cross-references: EMBL:J03487; NID:gl72111; PID:gl72112  
 C:Genetics:  
 A:Gene: SGD:PDR1  
 A:Cross-references: SGD:S0002981; MIPS:YGL013c  
 A:Map position: 7L  
 C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluste  
 C:Keywords: transmembrane protein  
 F:41-77/Domain: GAL4 zinc binuclear cluster homology <GAL4>  
 F:754-770/Domain: transmembrane #status predicted <TM1>  
 F:812-828/Domain: transmembrane #status predicted <TM2>

Query Match 57.4%; Score 35; DB 2; Length 1088;  
 Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
 | | | | |  
 DB 53 KIKCKGKFC 62

RESULT 15  
 S23862  
 hypothetical protein 1 - garlic mosaic virus  
 C:Species: garlic mosaic virus, GarMV  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Sep-1997  
 C:Accession: S23862  
 R:Choi, J.N.; Choi, Y.H.; Choi, Y.D.; Lee, J.S.  
 submitted to the EMBL Data Library, June 1992  
 A:Description: Nucleotide sequence of a cDNA for garlic mosaic virus.  
 A:Reference number: S23862  
 A:Accession: S23862

A:Status: Preliminary  
A:Molecule type: genomic RNA  
A:Residues: 1-345 <CHO>  
A:Cross-References: EMBL:X67134; NID:g59391; PID:g59392

Query Match 57.4%; Score 35; DB 2; Length 345;  
Best Local Similarity 57.1%; Pred. No. 57;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRFKFKC 10  
|:|:|:|  
Db 324 CEFRC 330

Search completed: September 7, 1999, 23:06:29  
Job time: 2477 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:17 ; Search time 71.87 seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-15  
Perfect score: 61  
Sequence: 1 KFKCKFKFKC 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	68.9	137	PA2M_AGGKCL	P49121 agkistrodon
2	39.5	64.8	1060	2FH1_DROME	P28166 drosophila
3	36	59.0	43	MYX1_CROCV	P12028 crotalus vi
4	36	59.0	42	MYX1_CROCV	P01476 crotalus vi
5	36	59.0	43	MYX2_CROCV	P12029 crotalus vi
6	36	59.0	43	MYX2_CROCV	P19861 crotalus vi
7	36	59.0	45	MYX2_CROCV	P01477 crotalus vi
8	36	59.0	121	PA2H_AGGKPI	P04361 agkistrodon
9	36	59.0	291	SUAR_RAT	P17988 rattus norv
10	36	59.0	398	YOFD_BACSU	P54469 bacillus su
11	35	57.4	2616	NDL_DROME	P98159 drosophila
12	35	57.4	1063	PDR1_YEAST	P12383 saccharomyc
13	34	55.7	395	NH10_CAEEL	P41999 caenorhabdi
14	34	55.7	373	REP1_YEAST	P03871 saccharomyc
15	34	55.7	263	VCP_VACCV	P10998 vaccinia vi
16	34	55.7	1361	YME9_YEAST	Q04693 saccharomyc
17	34	55.7	1246	YMY2_CAEEL	P34504 caenorhabdi
18	33	54.1	934	CO6_HUMAN	P13671 homo sapien
19	33	54.1	4655	LRP2_HUMAN	P98164 homo sapien
20	33	54.1	4660	LRP2_RAT	P98158 rattus norv
21	33	54.1	51	MYX4_CRODU	P24334 crotalus du
22	33	54.1	42	MYXC_CRODU	P01475 crotalus du
23	33	54.1	45	MYXC_CAEEL	P24330 crotalus ad
24	33	54.1	515	NH35_CAEEL	Q17771 caenorhabdi
25	33	54.1	876	RPOB_NPVAC	P41452 autographa
26	32	52.5	1290	BXB_CLOBO	P10844 clostridium
27	32	52.5	656	DNLT_HELPY	O25336 helicobacte
28	32	52.5	88	MAT8_MOUSE	O61835 mus musculu
29	32	52.5	3224	N358_HUMAN	P49792 homo sapien
30	32	52.5	138	PA2L_VIPAA	P17935 vipera ammo
31	32	52.5	3080	POLG_ZYMYC	P18479 z genome po
32	32	52.5	3083	POLG_ZYMYR	Q89330 z genome po
33	32	52.5	363	RTC1_SCHPO	Q09870 schizosacch
34	32	52.5	461	THB1_HUMAN	P10828 homo sapien
35	32	52.5	461	THB1_MOUSE	P37242 mus musculu
36	32	52.5	461	THB1_RAT	P18113 rattus norv
37	32	52.5	411	THB1_SHEEP	Q28571 ovis aries
38	32	52.5	369	THB1_XENLA	P18116 xenopus lae
39	32	52.5	476	THB2_HUMAN	P37243 homo sapien
40	32	52.5	475	THB2_MOUSE	P37244 mus musculu
41	32	52.5	514	THB2_RAT	P37826 rattus norv
42	32	52.5	373	THB5_XENLA	P18117 xenopus lae
43	32	52.5	373	THB6_XENLA	P18118 xenopus lae

ALIGNMENTS

RESULT 1	PA2M_AGGKCL	STANDARD;	PRT; 137 AA.
ID	PA2M_AGGKCL	STANDARD;	PRT; 137 AA.
AC	P49121;		
DT	01-FEB-1996 (REL. 33, CREATED)		
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)		
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)		
DE	PHOSPHOLIPASE A2 HOMOLOG, MYOTOXIN PRECURSOR.		
GN	MT1.		
OS	AGKISTRODON CONTORTRIX LATICINCTUS.		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;		
OC	SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;		
OC	AGKISTRODON.		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=VENOM GLAND;		
RX	MEDLINE; 96154243.		
RA	DE ARAUJO H.S., WHITE S.P., OWNBY C.L.;		
RT	"CDNA cloning and sequence analysis of a lysine-49 phospholipase A2		
RL	myotoxin from Agkistrodon contortrix laticinctus snake venom.";		
RL	ARCH. BIOCHEM. BIOPHYS. 326:21-30(1996)		
CC	-1- FUNCTION: THIS MYOTOXIC PROTEIN LACKS ENZYMATIC ACTIVITY. INDUCES		
CC	NECROSIS OF MUSCLE CELLS.		
CC	-1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; U21335; G809485; -		
DR	PROSITE; PS00118; PA2_HIS; 1.		
DR	PROSITE; PS00119; PA2_ASP; 1.		
DR	PFAM; PF00068; phoslip; 1.		
DR	HSSP; P04361; 1PPA.		
KW	TOXIN; VENOM; MULTIGENE FAMILY; SIGNAL.		
FT	SIGNAL 1 16 POTENTIAL.		
FT	CHAIN 17 137 MYOTOXIN.		
FT	ACT_SITE 63 63 BY SIMILARITY.		
FT	ACT_SITE 105 105 BY SIMILARITY.		
FT	DISULFID 42 131 BY SIMILARITY.		
FT	DISULFID 44 60 BY SIMILARITY.		
FT	DISULFID 59 111 BY SIMILARITY.		
FT	DISULFID 65 137 BY SIMILARITY.		
FT	DISULFID 66 104 BY SIMILARITY.		
FT	DISULFID 73 97 BY SIMILARITY.		
FT	DISULFID 91 102 BY SIMILARITY.		
SQ	SEQUENCE 137 AA; 15775 MW; 33BE1834 CRC32;		
Query Match	68.9%;	Score 42;	DB 1; Length 137;
Best Local Similarity	70.0%;	Pred. No. 1.3;	
Matches	7; Conservative	1; Mismatches	2; Indels 0; Gaps 0;
Qy	1 KFKCKFKFKC 10		
	1:1		
Db	122 KYKAVFKFKC 131		
RESULT 2	2FH1_DROME	STANDARD;	PRT; 1060 AA.
ID	2FH1_DROME	STANDARD;	PRT; 1060 AA.
AC	P28166;		

DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE ZINC-FINGER PROTEIN 1 (ZINC-FINGER HOMEODOMAIN PROTEIN 1).  
 GN ZFH-1.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92001539.  
 RA FORTINI M.E., LAI Z., RUBIN G.M.;  
 RA "The Drosophila zfh-1 and zfh-2 genes encode novel proteins  
 RT containing both zinc-finger and homeodomain motifs.";  
 RL MECH. DEV. 34:113-122(1991).  
 CC -!- FUNCTION: INVOLVED IN THE DEVELOPMENT OF THE EMBRYONIC CENTRAL  
 CC NERVOUS SYSTEM, EMBRYONIC MESODERM AND ADULT MUSCULATURE.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: MESODERM AND MESODERMALLY-DERIVED STRUCTURES  
 CC IN THE EMBRYO INCLUDING THE DORSAL VESSEL, SUPPORT CELLS OF THE  
 CC GONADS, AND SEGMENT-SPECIFIC ARRAYS OF ADULT MUSCLE PRECURSOR.  
 CC ALSO IDENTIFIED IN MOTOR NEURONS OF DEVELOPING CNS.  
 CC -!- SIMILARITY: TO HUMAN N1L-2-A.  
 CC -!- SIMILARITY: CONTAINS ONE HOMEODOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M63449; G158821; -.  
 DR PIR; S27816; S27816.  
 DR PIR; S33641; S33641.  
 DR FLYBASE; FBgn0004606; zfh1.  
 DR PROSITE; PS00027; HOMEODOMAIN\_1; 1.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 6.  
 DR PROSITE; PS00071; HOMEODOMAIN\_2; 1.  
 DR PFAM; PF00046; homeobox; 1.  
 DR PFAM; PF00096; zf-C2H2; 9.  
 DR HSSP; P08153; 1ZFD.  
 DR TRANSFAC; T00919; -.  
 KW ZINC-FINGER; METAL-BINDING; DNA-BINDING; HOMEODOMAIN; NUCLEAR PROTEIN.  
 FT ZN\_FING 74 97  
 FT DOMAIN 222 252  
 FT ZN\_FING 295 317  
 FT ZN\_FING 330 352  
 FT ZN\_FING 361 383  
 FT ZN\_FING 389 413  
 FT ZN\_FING 634 657  
 FT ZN\_FING 705 764  
 FT DNA\_BIND 973 995  
 FT ZN\_FING 1001 1023  
 FT ZN\_FING 1029 1050  
 FT ZN\_FING 1060 1060  
 SQ SEQUENCE 1060 AA; 117413 MW; 52AA59E8 CRC32;  
 Query Match 64.8%; Score 39.5; DB 1; Length 1060;  
 Best Local Similarity 64.3%; Pred. No. 17;  
 Matches 9; Conservative 0; Mismatches 0; Indels 5; Gaps 1;  
 QY 1 KFCKK-----FKFK 9  
 Db 360 KFCKCKCDKAFKFK 373  
 RESULT 3  
 MYX1\_CROVC  
 ID MYX1\_CROVC STANDARD; PRT; 43 AA.  
 AC P12028;

DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN I.  
 OS CROTALUS VIRIDIS CONCOLOR (MIDGET FADED RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCHEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 87320413.  
 RA BIEBER A.L., MCPARLAND R.H., BECKER R.R.;  
 RA "Amino acid sequences of myotoxins from Crotalus viridis concolor  
 RT venom.";  
 RL TOXICON 25:677-680(1987).  
 CC -!- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACTS EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -!- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 DR PIR; A29089; A29089.  
 DR PROSITE; PS00459; MYOTOXINS; 1.  
 DR PFAM; PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN.  
 FT DISULFID 4 36  
 FT DISULFID 11 30  
 FT DISULFID 18 37  
 FT DISULFID 43 AA; 5061 MW; 961CD130 CRC32;  
 SQ SEQUENCE 43 AA; 5061 MW; 961CD130 CRC32;  
 Query Match 59.0%; Score 36; DB 1; Length 43;  
 Best Local Similarity 50.0%; Pred. No. 4.6;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFCKKFKFKC 10  
 Db 27 KMDCRWKWC 36  
 RESULT 4  
 MYX1\_CROVC  
 ID MYX1\_CROVC STANDARD; PRT; 42 AA.  
 AC P01476;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN A (MYOTOXIN 1).  
 OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCHEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 79124714.  
 RA FOX J.W., ELZINGA M., TU A.T.;  
 RA "Amino acid sequence and disulfide bond assignment of myotoxin a  
 RT isolated from the venom of prairie rattlesnake (Crotalus viridis  
 RT viridis).";  
 RL BIOCHEMISTRY 18:678-684(1979).  
 RN [2]  
 RP SEQUENCE.  
 RC TISSUE=VENOM;  
 RX MEDLINE; 91263105.  
 RA AIRD S.D., KRUGGEL W.G., KAISER I.I.;  
 RA "Multiple myotoxin sequences from the venom of a single prairie  
 RT rattlesnake (Crotalus viridis viridis).";  
 RL TOXICON 29:265-268(1991).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE; 93037463.  
 RA BAKER B., UTAISINGHARON P., TU A.T.;  
 RT "Structure-function relationship of myotoxin a using peptide

fragments.";  
 RL ARCH. BIOCHEM. BIOPHYS. 298:325-331(1992).  
 CC -!- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACTS EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -!- LD(50) IS 3.0 MG/KG BY INTRAMUSCULAR INJECTION.  
 CC -!- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 DR PIR; A01736; MKRSNV.  
 DR PIR; S27115; S27115.  
 DR PROSITE; PS00459; MYOTOXINS; 1.  
 DR PFAM; PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN. 4 36  
 FT DISULFID 11 30  
 FT DISULFID 18 37  
 FT VARIANT 19 19  
 FT VARIANT 25 25  
 FT VARIANT 25 25  
 FT VARIANT 25 25  
 FT VARIANT 25 25  
 FT VARIANT 33 33  
 SQ SEQUENCE 42 AA; 4828 MW; 624F184E CRC32;  
 Query Match 59.0%; Score 36; DB 1; Length 42;  
 Best Local Similarity 50.0%; Pred. No. 4.5;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKCKEKFCK 10  
 Db 27 KMDCRWKWC 36  
 RESULT 5  
 MYX2\_CROV  
 ID MYX2\_CROV STANDARD; PRT; 43 AA.  
 AC P12029;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN II.  
 OS CROTALUS VIRIDIS CONCOLOR (MIDGET FADED RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 87320413.  
 RA BIEBER A.L., MCPARLAND R.H., BECKER R.R.;  
 RT "Amino acid sequences of myotoxins from Crotalus viridis concolor  
 venom.";  
 RL TOXICON 25:677-680(1987).  
 CC -!- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACTS EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -!- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 DR PIR; B29089; B29089.  
 DR PROSITE; PS00459; MYOTOXINS; 1.  
 DR PFAM; PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN. 4 36  
 FT DISULFID 11 30  
 FT DISULFID 18 37  
 FT VARIANT 15 15  
 FT VARIANT 15 15  
 FT VARIANT 19 19  
 FT VARIANT 33 33  
 SQ SEQUENCE 43 AA; 5034 MW; E7CDAEF2 CRC32;  
 Query Match 59.0%; Score 36; DB 1; Length 42;  
 Best Local Similarity 50.0%; Pred. No. 4.5;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKCKEKFCK 10  
 Db 27 KMDCRWKWC 36  
 RESULT 5  
 MYX2\_CROV  
 ID MYX2\_CROV STANDARD; PRT; 43 AA.  
 AC P12029;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN II.  
 OS CROTALUS VIRIDIS CONCOLOR (MIDGET FADED RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 87320413.  
 RA BIEBER A.L., MCPARLAND R.H., BECKER R.R.;  
 RT "Amino acid sequences of myotoxins from Crotalus viridis concolor  
 venom.";  
 RL TOXICON 25:677-680(1987).  
 CC -!- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACTS EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -!- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 DR PIR; B29089; B29089.  
 DR PROSITE; PS00459; MYOTOXINS; 1.  
 DR PFAM; PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN. 4 36  
 FT DISULFID 11 30  
 FT DISULFID 18 37  
 FT VARIANT 15 15  
 FT VARIANT 15 15  
 FT VARIANT 19 19  
 FT VARIANT 33 33  
 SQ SEQUENCE 43 AA; 5034 MW; E7CDAEF2 CRC32;  
 Query Match 59.0%; Score 36; DB 1; Length 45;  
 Best Local Similarity 50.0%; Pred. No. 4.7;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKCKEKFCK 10  
 Db 27 KMDCRWKWC 36  
 RESULT 7  
 MYXC\_CROVH  
 ID MYXC\_CROVH STANDARD; PRT; 43 AA.  
 AC P01477;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN (TOXIC PEPTIDE C).  
 OS CROTALUS VIRIDIS HELLERI (SOUTHERN PACIFIC RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 79015339.

Query Match 59.0%; Score 36; DB 1; Length 43;  
 Best Local Similarity 50.0%; Pred. No. 4.6;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKCKEKFCK 10  
 Db 27 KMDCRWKWC 36  
 RESULT 6  
 MYX2\_CROV  
 ID MYX2\_CROV STANDARD; PRT; 45 AA.  
 AC P19861;  
 DT 01-FEB-1991 (REL. 17, CREATED)  
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXINS 2 AND 3.  
 OS CROTALUS VIRIDIS VIRIDIS (PRAIRIE RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX TISSUE-VENOM;  
 RX MEDLINE; 91071443.  
 RA GRIFFIN P.R., AIRD S.D.;  
 RT "A new small myotoxin from the venom of the prairie rattlesnake  
 (Crotalus viridis viridis).";  
 RL FEBS LETT. 274:43-47(1990).  
 CC -!- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACTS EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -!- THE SEQUENCE SHOWN IS THAT OF MYOTOXIN 2.  
 CC -!- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 DR PIR; S12909; S12909.  
 DR PROSITE; PS00459; MYOTOXINS; 1.  
 DR PFAM; PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN. 4 36  
 FT DISULFID 11 30  
 FT DISULFID 18 37  
 FT VARIANT 8 8  
 SQ SEQUENCE 45 AA; 5246 MW; 41B35120 CRC32;  
 Query Match 59.0%; Score 36; DB 1; Length 45;  
 Best Local Similarity 50.0%; Pred. No. 4.7;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KFKCKEKFCK 10  
 Db 27 KMDCRWKWC 36  
 RESULT 7  
 MYXC\_CROVH  
 ID MYXC\_CROVH STANDARD; PRT; 43 AA.  
 AC P01477;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MYOTOXIN (TOXIC PEPTIDE C).  
 OS CROTALUS VIRIDIS HELLERI (SOUTHERN PACIFIC RATTLESNAKE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC CROTALUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 79015339.

RA MAEDA N., TAMIYA N., PATTABHIRAMAN T.R., RUSSELL F.E.;  
 RT "Some chemical properties of the venom of the rattlesnake, Crotalus  
 RL viridis helleri.";  
 CC TOXICON 16:431-441(1978).  
 CC -1- FUNCTION: CAUSES SEVERE MUSCLE NECROSIS BY A NON-ENZYMATIC  
 CC MECHANISM. ACSX EXTREMELY RAPIDLY AND SERVES TWO PRIMARY  
 CC FUNCTIONS: LIMIT THE FLIGHT OF PREY BY CAUSING INSTANTANEOUS  
 CC PARALYSIS OF THE HIND LIMBS AND PROMOTE RAPID DEATH BY PARALYSIS  
 CC OF THE DIAPHRAGM.  
 CC -1- LD(50) IS 1.96 MG/KG BY INTRAVENOUS INJECTION.  
 CC -1- SIMILARITY: BELONGS TO THE MYOTOXIN FAMILY OF SNAKE TOXINS.  
 CC PIR: A01737; CXRSCH.  
 DR PROSITE: PS00459; MYOTOXINS; 1.  
 DR PFAM: PF00819; Myotoxins; 1.  
 KW VENOM; TOXIN.  
 FT DISULFID 4 36 BY SIMILARITY.  
 FT DISULFID 11 30 BY SIMILARITY.  
 FT DISULFID 18 37 BY SIMILARITY.  
 SQ SEQUENCE 43 AA; 4989 MW; 678C7E5B CRC32;

Query Match 59.0%; Score 36; DB 1; Length 43;  
 Best Local Similarity 50.0%; Pred. No. 4.6;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFCKKFKFC 10  
 | | | | |  
 Db 27 KMDCRWKWC 36

RESULT 8  
 PA2H\_AKGPI STANDARD; PRT; 121 AA.  
 AC P04361;  
 DT 20-MAR-1987 (REL. 04, CREATED)  
 DT 20-MAR-1987 (REL. 04, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PHOSPHOLIPASE A2 HOMOLOG (APP-K-49).  
 OS AGKISTRODON PISCIVORUS PISCIVORUS (EASTERN COTTONMOUTH).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SOUAMATA;  
 OC SCLEROGLOSSA; SERPENTES; COLUBROIDEA; VIPERIDAE; CROTALINAE;  
 OC AGKISTRODON.  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=VENOM;  
 RX MEDLINE; 86168190.  
 RA MARAGANORE J.M., HEINRIKSON R.L.;  
 RT "The lysine-49 phospholipase A2 from the venom of Agkistrodon  
 RT piscivorus piscivorus. Relation of structure and function to other  
 RT phospholipases A2.";  
 RL J. BIOL. CHEM. 261:4797-4804(1986).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE; 91009222.  
 RA HOLLAND D.R., CLANCY L.L., MUCHMORE S.W., RYDE T.J., EINSIPAHR H.M.,  
 RA FINZEL B.C., HEINRIKSON R.L., WATENPAUGH K.D.;  
 RT "The crystal structure of a lysine 49 phospholipase A2 from the venom  
 RT of the cottonmouth snake at 2.0-A resolution.";  
 RL J. BIOL. CHEM. 265:17649-17656(1990).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE; 93054566.  
 RA SCOTT D.L., ACHARI A., VIDAL J.C., SIGLER P.B.;  
 RT "Crystallographic and biochemical studies of the (inactive) Lys-49  
 RT phospholipase A2 from the venom of Agkistrodon piscivorus  
 RT piscivorus.";  
 RL J. BIOL. CHEM. 267:22645-22657(1992).  
 CC -1- FUNCTION: THIS PROTEIN HAS NO PHOSPHOLIPASE A2 ACTIVITY. IT IS  
 CC AN EDEMA-INDUCING PA2.  
 CC -1- COFACTOR: DOES NOT REQUIRE CALCIUM. IT HAS LOST ONE OF THE  
 CC CALCIUM BINDING SITE (ASP->LYS IN POSITION 48).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- LD(50) IS 25 MG/KG BY INTRAVENOUS INJECTION.

CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 DR PIR: A00767; PSSNAM.  
 DR PDB: 1PPA; 31-JAN-94.  
 DR PROSITE: PS00118; PA2\_HIS; 1.  
 DR PROSITE: PS00119; PA2\_ASP; 1.  
 DR PFAM: PF00068; Phoslip; 1.  
 KW MULTIGENE FAMILY; VENOM; 3D-STRUCTURE.  
 FT ACT\_SITE 47 47 BY SIMILARITY.  
 FT ACT\_SITE 89 89 BY SIMILARITY.  
 FT DISULFID 26 115  
 FT DISULFID 28 44  
 FT DISULFID 43 95  
 FT DISULFID 49 121  
 FT DISULFID 50 88  
 FT DISULFID 57 81  
 FT DISULFID 75 86  
 FT HELIX 2 13  
 FT HELIX 17 21  
 FT TURN 22 23  
 FT TURN 25 28  
 FT TURN 30 31  
 FT TURN 39 51  
 FT TURN 52 53  
 FT TURN 59 61  
 FT STRAND 67 68  
 FT STRAND 73 74  
 FT HELIX 80 98  
 FT TURN 99 99  
 FT HELIX 100 102  
 FT HELIX 105 107  
 SQ SEQUENCE 121 AA; 13961 MW; C13E7B5C CRC32;

Query Match 59.0%; Score 36; DB 1; Length 121;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFCKKFKFC 10  
 | | | | |  
 Db 106 KYKAYFKLC 115

RESULT 9  
 SUAR\_RAT STANDARD; PRT; 291 AA.  
 ID SUAR\_RAT  
 AC P17988;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE ARYL SULFOTRANSFERASE (EC 2.8.2.1) (PHENOL SULFOTRANSFERASE) (PST-1)  
 DE (SULFOKINASE) (ARYL SULFOTRANSFERASE IV) (ASTIV) (TYROSINE-ESTER  
 DE SULFOTRANSFERASE) (MINOXIDIL SULFOTRANSFERASE).  
 GN ST1AL.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RX MEDLINE; 90326537.  
 RA OZAWA S., NAGATA K., GONG D., YAMAZOE Y., KATO R.;  
 RT "Nucleotide sequence of a full-length cDNA (PST-1) for aryl  
 RT sulfotransferase from rat liver.";  
 RL NUCLEIC ACIDS RES. 18:4001-4001(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RX MEDLINE; 94131305.  
 RA KHAN A.S., TAYLOR B.R., CHUNG K., ETHEREDGE J., GONZALES R.,  
 RA RINGER D.P.;  
 RT "Genomic structure of rat liver aryl sulfotransferase IV-encoding  
 RT gene.";  
 RL GENE 137:321-326(1993).



RN [3]  
RP SEQUENCE OF 7-291 FROM N.A.  
RX MEDLINE: 92379794.  
RA YEROKUN T., ETHEREDGE J.L., NORTON T.R., CARTER H.A., CHUNG K.H.,  
RA BIRCKBICHLER P.J., RINGER D.P.;  
RT "Characterization of a complementary DNA for rat liver aryl  
RT sulfotransferase IV and use in evaluating the hepatic gene transcript  
RT levels of rats at various stages of 2-acetylaminofluorene-induced  
RT hepatocarcinogenesis.";  
RL CANCER RES. 52:4779-4786(1992).  
RN [4]  
RP CHARACTERIZATION.  
RP TISSUE=LIVER.  
RX MEDLINE: 93191709.  
RA CRUICKSHANK D., SANSOM L.N., VERONESE M.E., MOJARRABI B.,  
RA MCMAHON M.E., ZHU X.;  
RT "cDNA expression studies of rat liver aryl sulphotransferase.";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 191:295-301(1993).  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE: 95195680.  
RA YAMAZOE Y., OZAWA S., NAGATA K., GONG D.-W., KATO R.;  
RT "Characterization and expression of hepatic sulfotransferase involved  
RT in the metabolism of N-substituted aryl compounds.";  
RL ENVIRON. HEALTH PERSPECT. 102:99-103(1994).  
RN [6]  
RP SEQUENCE OF 63-68, AND CHARACTERIZATION.  
RX MEDLINE: 95074030.  
RA ZHENG Y., BERGOLD A., DUFFEL M.W.;  
RT "Affinity labeling of aryl sulfotransferase IV. Identification of a  
RT peptide sequence at the binding site for 3'-phosphoadenosine-5'-  
RT phosphosulfate.";  
RL J. BIOL. CHEM. 269:30313-30319(1994).  
RN [7]  
RP CHARACTERIZATION.  
RX MEDLINE: 94306583.  
RA RUNGE-MORRIS M.A.;  
RT "Sulfotransferase gene expression in rat hepatic and extrahepatic  
RT tissues.";  
RL CHEM. BIOL. INTERACT. 92:67-76(1994).  
CC -!- FUNCTION: CATALYZES O-SULFATION OF PHENOLS, N-O- SULFATION OF  
CC MINOXIDIL AND TYROSINE ESTERS  
CC -!- CATALYTIC ACTIVITY: 3'-PHOSPHOADENYLSULFATE + A PHENOL -  
CC ADENOSINE 3',5'-BISPHOSPHATE + AN ARYL SULFATE.  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- TISSUE SPECIFICITY: LIVER, KIDNEY, HEART AND COLON.  
CC -!- INDUCTION: INDUCED BY ANDROGENS AND SUPPRESSED BY ESTROGENS. THE  
CC EXPRESSION IS UNDER THE INFLUENCE OF PITUITARY GROWTH HORMONE AND  
CC THYROID HORMONE.  
CC -!- PTM: THE N-TERMINAL IS BLOCKED.  
CC -!- SIMILARITY: BELONGS TO THE SULFOTRANSFERASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L19998; G310179; -  
DR EMBL; X52883; G55761; -  
DR EMBL; L16241; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X68640; G55765; -  
DR PIR; S10329; S10329.  
KW PFAM; PF00685; Sulfotransferase; 1.  
FT TRANSFERASE; STEROID METABOLISM.  
FT BINDING 255 261 PAPS BINDING SITE (POTENTIAL).  
SQ SEQUENCE 291 AA; 33906 MW; A6635E56 CRC32;

Query Match

59.0%; Score 36; DB 1; Length 291;

Best Local Similarity 71.4%; Pred. No. 22;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 CKFKFKC 10  
DB 283 CDFKFKC 289  
| | | | |  
RESULT 10  
YQFD\_BAGSU STANDARD; PRT; 398 AA.  
ID YQFD\_BAGSU STANDARD; PRT; 398 AA.  
AC P54469;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 45.7 KD PROTEIN IN RPSU-PHOH INTERGENIC REGION.  
GN YQFD.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
CC [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEMARU K., HOSONO S.,  
RA SATO T., TAKEUCHI M.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: TO B. MEGATERIUM SPOIV.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D84432; D1013155; -  
DR EMBL; Z99117; E1183765; -  
DR SUBTILIST; BGL1654; YQFD.  
KW HYPOTHETICAL PROTEIN; SPOULATION.  
SQ SEQUENCE 398 AA; 45727 MW; 50217302 CRC32;  
Query Match 59.0%; Score 36; DB 1; Length 398;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFKCKFKC 8  
DB 64 KFKCKARF 71  
| | | | |  
RESULT 11  
NDL\_DROME STANDARD; PRT; 2616 AA.  
ID NDL\_DROME STANDARD; PRT; 2616 AA.  
AC P98159;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE SERINE PROTEASE NUDEL PRECURSOR (EC 3.4.21.-).  
GN NDL.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
CC [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON-R; TISSUE=OVARY;  
RX MEDLINE: 95401268  
RA HONG C.C., HASHIMOTO C.;  
RT "An unusual mosaic protein with a protease domain, encoded by the  
RT ndel gene, is involved in defining embryonic dorsoventral polarity  
RT in Drosophila.";  
RT

RL CELL 82:785-794(1995).

CC -!- FUNCTION: NUDEL, PIPE AND WINDBEUTEL TOGETHER TRIGGER THE PROTEASE

CC CASCADE WITHIN THE EXTRAEMBRYONIC PERIVITELLINE COMPARTMENT WHICH

CC INDUCES DORSOVENTRAL POLARITY OF THE DROSOPHILA EMBRYO. NUDEL IS

CC DIRECTLY INVOLVED IN LOCALLY PRODUCING THE TOLL LIGAND.

CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.

CC -!- TISSUE SPECIFICITY: FOLLICLE.

CC -!- PTM: REQUIRES CLEAVAGE FOR ACTIVATION (PRESUMABLY).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

CC -!- SIMILARITY: CONTAINS 11 LDL-RECEPTOR CLASS A DOMAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U29153; G984321; ..

DR FLYBASE: FBgn0002926; ndi.

DR PROSITE: PS00134; TRYPSIN\_HIS; 1.

DR PROSITE: PS00135; TRYPSIN\_SER; 1.

DR PROSITE: PS01209; LDLRA\_1; 6.

DR PROSITE: PS00068; LDLRA\_2; 8.

DR PFAM: PF00057; ldl\_recept\_a; 7.

DR PFAM: PF00089; trypsin; 1.

DR HSSP: P00763; lDPO.

DR SERINE PROTEASE: GLYCOPROTEIN; SIGNAL; DEVELOPMENTAL PROTEIN;

KW HYDROLASE; REPEAT; ZYMOGEN; EXTRACELLULAR MATRIX.

FT SIGNAL 1 43 POTENTIAL.

FT CHAIN 44 2616 SERINE PROTEASE NUDEL.

FT REPEAT 261 269 WIID 1.

FT REPEAT 320 328 WIID 2.

FT REPEAT 399 407 WIID 3.

FT REPEAT 446 454 WIID 4.

FT REPEAT 477 485 WIID 5.

FT REPEAT 528 536 WIID 6.

FT DOMAIN 889 929 LDL-RECEPTOR CLASS A 1.

FT DOMAIN 929 956 LDL-RECEPTOR CLASS A 2 (PARTIAL).

FT DOMAIN 955 1006 LDL-RECEPTOR CLASS A 3.

FT DOMAIN 1144 1379 CATALYTIC.

FT DOMAIN 1394 1432 SER/THR-RICH.

FT DOMAIN 1489 1702 LDL-RECEPTOR CLASS A 5 (PARTIAL).

FT DOMAIN 1713 1743 LDL-RECEPTOR CLASS A 6 (PARTIAL).

FT DOMAIN 1745 1775 LDL-RECEPTOR CLASS A 7.

FT DOMAIN 1774 1813 LDL-RECEPTOR CLASS A 8.

FT DOMAIN 2039 2138 C-TERMINAL SERINE PROTEASE-LIKE.

FT DOMAIN 2308 2346 LDL-RECEPTOR CLASS A 9.

FT DOMAIN 2349 2389 LDL-RECEPTOR CLASS A 10 (PARTIAL).

FT DOMAIN 2387 2419 LDL-RECEPTOR CLASS A 11.

FT SITE 1031 1033 CELL ATTACHMENT SITE (POTENTIAL).

FT ACT\_SITE 1185 1185 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 1233 1233 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 1332 1332 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 891 905 BY SIMILARITY.

FT DISULFID 899 918 BY SIMILARITY.

FT DISULFID 912 927 BY SIMILARITY.

FT DISULFID 957 982 BY SIMILARITY.

FT DISULFID 964 995 BY SIMILARITY.

FT DISULFID 989 1004 BY SIMILARITY.

FT DISULFID 1170 1186 BY SIMILARITY.

FT DISULFID 1276 1338 POTENTIAL.

FT DISULFID 1305 1317 BY SIMILARITY.

FT DISULFID 1328 1359 BY SIMILARITY.

FT DISULFID 1396 1408 BY SIMILARITY.

FT DISULFID 1401 1421 BY SIMILARITY.

FT DISULFID 1415 1430 BY SIMILARITY.

FT DISULFID 1776 1789 BY SIMILARITY.

FT DISULFID 1783 1802 BY SIMILARITY.

FT DISULFID 1796 1811 BY SIMILARITY.

FT DISULFID 2310 2320 BY SIMILARITY.

FT DISULFID 2315 2333 BY SIMILARITY.

FT DISULFID 2327 2344 BY SIMILARITY.

FT DISULFID 2351 2364 BY SIMILARITY.

FT DISULFID 2358 2377 BY SIMILARITY.

FT DISULFID 2371 2387 BY SIMILARITY.

FT DISULFID 2421 2435 BY SIMILARITY.

FT DISULFID 2428 2448 BY SIMILARITY.

FT DISULFID 2442 2457 BY SIMILARITY.

FT CARBOHYD 291 291 POTENTIAL.

FT CARBOHYD 347 347 POTENTIAL.

FT CARBOHYD 379 379 POTENTIAL.

FT CARBOHYD 417 417 POTENTIAL.

FT CARBOHYD 492 492 POTENTIAL.

FT CARBOHYD 515 515 POTENTIAL.

FT CARBOHYD 598 598 POTENTIAL.

FT CARBOHYD 794 794 GLYCOSAMINOGLYCAN (POTENTIAL).

FT CARBOHYD 827 827 POTENTIAL.

FT CARBOHYD 829 829 GLYCOSAMINOGLYCAN (POTENTIAL).

FT CARBOHYD 861 861 POTENTIAL.

FT CARBOHYD 975 975 POTENTIAL.

FT CARBOHYD 1064 1064 POTENTIAL.

FT CARBOHYD 1445 1445 POTENTIAL.

FT CARBOHYD 1878 1878 POTENTIAL.

FT CARBOHYD 1956 1956 POTENTIAL.

FT CARBOHYD 2023 2023 POTENTIAL.

FT CARBOHYD 2144 2144 POTENTIAL.

FT CARBOHYD 2173 2173 POTENTIAL.

FT CARBOHYD 2197 2197 POTENTIAL.

FT CARBOHYD 2237 2237 POTENTIAL.

FT CARBOHYD 2269 2269 POTENTIAL.

FT CARBOHYD 2420 2420 POTENTIAL.

FT CARBOHYD 2556 2556 POTENTIAL.

FT CARBOHYD 2601 2601 POTENTIAL.

SQ SEQUENCE 2616 AA; 292371 MW; OFFD8412 CRC32;

Query Match 57.4%; Score 35; DB 1; Length 2616;

Best Local Similarity 50.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFKC 10

DB 2355 EMKRTSFKC 2364

RESULT 12

ID PDRL\_YEAST STANDARD; PRT; 1063 AA.

AC P12383;

DT 01-OCT-1989 (REL. 12, CREATED)

DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE PLEIOTROPIC DRUG RESISTANCE REGULATORY PROTEIN 1.

GN PDRI OR ANTI OR BOR2 OR CYH3 OR NRA2 OR SMR2 OR YGL013C.

OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;

OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 88059084.

RT BALZI E., CHEN W., ULASZEWSKI S., CAPIEAUX E., GOFFEAU A.;

RL "The multidrug resistance gene PDRI from *Saccharomyces cerevisiae*.";

RN J. BIOL. CHEM. 262:16871-16879(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91353083.

RA CHEN W., BALZI E., CAPIEAUX E., CHODER M., GOFFEAU A.;

RT "The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1

RT and ATEL loci on chromosome VII from *Saccharomyces cerevisiae* reveals

RT the PDRI gene, a new member of the genetic network controlling

RT pleiotropic drug resistance.";

RL YEAST 7:287-299(1991).

RN SEQUENCE FROM N.A.  
 RA HEBLING U., HOFMANN B., DELIUS H.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: POSITIVE REGULATOR OF PROTEINS INVOLVED IN PERMEABILITY.  
 CC PDR1 AND PDR3 JOINTLY CONTROL THE TRANSCRIPTION LEVEL OF BOTH SNQ2  
 CC AND PDR5.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR  
 CC CLUSTER DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03487; G172112; -.  
 DR EMBL; S58126; E9139; ALT\_TERM.  
 DR EMBL; 272535; E243845; -.  
 DR PIR; A28466; A28466.  
 DR PIR; S31560; S31560.  
 DR SGD; L0001361; PDR1.  
 DR PROSITE; PS00463; ZN2\_CYS6\_FUNGAL\_1; 1.  
 DR PROSITE; PS00468; ZN2\_CYS6\_FUNGAL\_2; 1.  
 DR PFAM; PF00172; Zn\_c1us; 1.  
 DR HSP; P08657; 1CLD.  
 KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC; METAL-BINDING.  
 FT DNA\_BIND 46 72 ZN(2)-CYS(6), FUNGAL-TYPE.  
 FT CONFLICT 411 411 K -> Q (IN REF. 3).  
 FT CONFLICT 530 530 A -> R (IN REF. 2).  
 FT CONFLICT 820 820 A -> T (IN REF. 3).  
 FT CONFLICT 921 921 I -> T (IN REF. 3).  
 FT CONFLICT 981 981 S -> T (IN REF. 3).  
 FT CONFLICT 1014 1014 N -> NNNNNN (IN REF. 3).  
 FT CONFLICT 1063 1063 N -> F06505F0 CRC32;  
 SQ SEQUENCE 1063 AA; 121190 MW; F06505F0 CRC32;  
  
 Query Match 57.4%; Score 35; DB 1; Length 1063;  
 Best Local Similarity 60.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 1 KFKCKFKFC 10  
 DB 53 KIKCKGKFC 62  
 ||| |||  
  
 RESULT 13  
 ID NH10 CAEEL STANDARD; PRT; 395 AA.  
 AC P41999;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE STEROID HORMONE RECEPTOR FAMILY MEMBER NHR-10.  
 GN NHR-10 OR B0280.8  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA FULTON L., WATERSTON R.;  
 RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U10438; G500769; -.  
 DR WORMPEP; B0280.8; C800815.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR PFAM; PF00104; hormone\_rec; 1.  
 DR PFAM; PF00105; zf-C4; 1.  
 DR HSP; P03372; 1HCO.  
 KW RECEPTOR; TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KW ZINC-FINGER.  
 FT DNA\_BIND 18 83 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 18 38 C4-TYPE.  
 FT ZN\_FING 54 78 C4-TYPE.  
 SQ SEQUENCE 395 AA; 44899 MW; 1B5CE2BC CRC32;  
  
 Query Match 55.7%; Score 34; DB 1; Length 395;  
 Best Local Similarity 55.6%; Pred. No. 57;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 FKCKFKFC 10  
 DB 52 FSCQFGKC 60  
 | | | | |  
  
 RESULT 14  
 ID REPI\_YEAST STANDARD; PRT; 373 AA.  
 AC P03871;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE TRANS-ACTING FACTOR B (REPI) (PROTEIN BAKER).  
 GN REPI.  
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC PLASMID 2-MICRON.  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
 OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A364A D5;  
 RX MEDLINE; 81012161.  
 RA HARTLEY J.L., DONELSON J.E.;  
 RT "Nucleotide sequence of the yeast plasmid."  
 RL NATURE 286:860-864(1980).  
 RN [2]  
 RP SEQUENCE OF 330-373 FROM N.A.  
 RX MEDLINE; 80034481.  
 RA HINDLEY J., PHEAR G.A.;  
 RT "Sequence of 1019 nucleotides encompassing one of the inverted  
 RT repeats from the yeast 2 micrometer plasmid."  
 RL NUCLEIC ACIDS RES. 7:361-375(1979).  
 CC -!- FUNCTION: PLASMID PARTITION REQUIRE REPI, REP2, AND A CIS-ACTING  
 CC DNA SEQUENCE (KNOWN AS STB). REP 1 MAY ACT BY INTERCALATING IN  
 CC THE YEAST NUCLEAR MATRIX AND BINDING STB EITHER DIRECTLY OR VIA  
 CC REP2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J01347; G172192; -.  
 DR EMBL; V01323; G4183; -.  
 DR PIR; A04503; PDBYB.  
 DR SGD; L0001611; REPI.  
 KW PLASMID; TRANS-ACTING FACTOR.

```

SQ SEQUENCE 373 AA: 43231 MW: 2B7C242B CRC32;
Query Match 55.7%; Score 34; DB 1; Length 373;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFKCKEK 7
   :|:|:|:|
Db 361 EFRCKEK 367

RESULT 15
VCP_VACCV STANDARD; PRT; 263 AA.
AC P10998;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE COMPLEMENT CONTROL PROTEIN PRECURSOR (VCP) (SECRETORY PROTEIN 35)
DE (PROTEIN C3) (28 KD PROTEIN).
GN C3L
OS VACCINIA VIRUS (STRAIN WR), AND VACCINIA VIRUS (STRAIN COPENHAGEN).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUS.
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 20-37.
RP STRAIN=WR;
RX MEDLINE; 88318974.
RA KOTWAL G.J., MOSS B.;
RT "Vaccinia virus encodes a secretory polypeptide structurally related
RL NATURE 335:176-178(1988).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=WR;
RX MEDLINE; 89073756.
RA KOTWAL G.J., MOSS B.;
RT "Analysis of a large cluster of nonessential genes deleted from a
RL vaccinia virus terminal transposition mutant.";
RL VIROLOGY 167:524-537(1988).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=COPENHAGEN;
RX MEDLINE; 91021027.
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RT "The complete DNA sequence of vaccinia virus.";
RL VIROLOGY 179:247-266(1990).
[4]
RN COMPLETE GENOME.
RP STRAIN=COPENHAGEN;
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:517-563(1990).
[5]
RN FUNCTION.
RP MEDLINE; 92115714.
RA ISAACS S.N., KOTWAL G.J., MOSS B.;
RT "Vaccinia virus complement-control protein prevents
RL antibody-dependent complement-enhanced neutralization of infectivity
RL and contributes to virulence.";
RL PROC. NATL. ACAD. SCI. U.S.A. 89:628-632(1992).
[6]
RN STRUCTURE BY NMR OF 146-263.
RX MEDLINE; 97446168.
RA WILES A.P., SHAW G., BRIGHT J., PERCZEL A., CAMPBELL I.D.,
RA BARLOW P.N.;
RT "NMR studies of a viral protein that mimics the regulators of
RL complement activation.";
RL J. MOL. BIOL. 272:253-265(1997).
CC -!- FUNCTION: SERVES TO PROTECT THE VIRUS AGAINST COMPLEMENT ATTACK BY
INHIBITING BOTH CLASSICAL AND ALTERNATIVE PATHWAYS OF COMPLEMENT
```

```

CC ACTIVATION. BINDS C3B AND C4B.
CC -!- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
CC COMPLEMENT ACTIVATION (RCA).
CC -!- SIMILARITY: CONTAINS 4 SUSHI (SCR) REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

DR EMBL; X13166; G60691;
DR EMBL; M2812; G893314;
DR EMBL; M35027; G335345;
DR PIR; A31005; WNVZSP.
DR PDB; 1VVC; 03-DEC-97.
DR PDB; 1VVE; 03-DEC-97.
DR PFAM; PF00084; sushi; 4.
KW SIGNAL; REPEAT; SUSHI; 3D-STRUCTURE.
FT SIGNAL 1 19
FT CHAIN 20 263
FT DOMAIN 20 262
FT REPEAT 20 82
FT REPEAT 85 144
FT REPEAT 147 202
FT REPEAT 205 262
FT DISULFID 21 70
FT DISULFID 54 81
FT DISULFID 86 126
FT DISULFID 112 143
FT DISULFID 148 190
FT DISULFID 176 201
FT DISULFID 206 248
FT DISULFID 234 261
SQ SEQUENCE 263 AA; 28629 MW; 4F70DDFE CRC32;
COMPLEMENT CONTROL PROTEIN.
4 X SUSHI (SCR) REPEATS.
SUSHI 1.
SUSHI 2.
SUSHI 3.
SUSHI 4.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
```

```

Query Match 55.7%; Score 34; DB 1; Length 263;
Best Local Similarity 62.5%; Pred. No. 41;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKCKEKFK 9
   |||:|:|
Db 232 FKCKYGYK 239
```

Search completed: September 7, 1999, 23:50:18  
Job time: 1957 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:34:46 ; Search time 148.39 Seconds  
(without alignments)  
4.147 Million cell updates/sec

Title: US-09-124-280A-15

Perfect score: 61

Sequence: 1 KCKFKFKC 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	68.9	441	4 O00149	O00149 homo sapien
2	42	68.9	517	4 O75530	O75530 homo sapien
3	42	68.9	535	11 P97462	P97462 mus musculus
4	41	67.2	170	1 O95556	O95556 pyrococcus
5	39	63.9	620	5 Q19807	Q19807 caenorhabdi
6	39	63.9	66	12 Q65257	Q65257 african swi
7	37	60.7	514	5 Q10956	Q10956 caenorhabdi
8	37	60.7	224	5 P91567	P91567 caenorhabdi
9	37	60.7	548	5 Q19496	Q19496 caenorhabdi
10	37	60.7	338	5 Q17928	Q17928 caenorhabdi
11	37	60.7	331	5 O96867	O96867 strongyloce
12	37	60.7	437	12 Q9YW73	Q9YW73 melanoplus
13	36	59.0	363	1 Q30192	Q30192 archaeglob
14	36	59.0	805	5 O09577	O09577 caenorhabdi
15	35	57.4	764	3 Q05712	Q05712 saccharomyc
16	35	57.4	652	4 O00274	O00274 homo sapien
17	35	57.4	347	5 Q17135	Q17135 caenorhabdi
18	35	57.4	477	5 O45666	O45666 caenorhabdi
19	35	57.4	971	5 O45785	O45785 caenorhabdi
20	35	57.4	142	5 O97146	O97146 sarcophaga
21	35	57.4	644	11 O89103	O89103 mus musculus
22	35	57.4	345	12 Q67663	Q67663 garlic mosa
23	35	57.4	1924	12 Q67654	Q67654 garlic late
24	34	55.7	380	3 Q07007	Q07007 saccharomyc
25	34	55.7	382	3 Q04137	Q04137 saccharomyc
26	34	55.7	607	5 P91452	P91452 caenorhabdi
27	34	55.7	572	5 Q23498	Q23498 caenorhabdi
28	34	55.7	2496	5 O96175	O96175 plasmodium
29	34	55.7	1817	5 O96206	O96206 plasmodium

#### ALIGNMENTS

##### RESULT 1

O00149 PRELIMINARY; PRT; 441 AA.

ID O00149

AC O00149;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE EMBRYONIC ECTODERM DEVELOPMENT PROTEIN HOMOLOG.

GN EED.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=SPLEEN, THYMUS, TESTIS, UTERUS, PROSTATE, SMALL INTESTINE,

RC COLON, PERIPHERAL BLOOD LYMPHOCYTES;

RA PEYTAI R., HONG S.-S., GAY B., DUPUY D'ANGEAC A., SELIG L., . . .

RA BENICHOUS S., BENAROUS R., BOULANGER P.;

RT "Heed, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1."

RL J. Biol. Chem. 0:0-0(1998).

DR EMBL; U90851; AAD08714.1; .

DR PFAM; PF00400; WD40; 2.

SQ SEQUENCE 441 AA; 50155 MW; 33EF5C6D CRC32;

Query Match 68.9%; Score 42; DB 4; Length 441;  
Best Local Similarity 75.0%; Pred. No. 4.6;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 3 KCKFKFKC 10  
Db 77 KCKYFKC 84

##### RESULT 2

O75530 PRELIMINARY; PRT; 517 AA.

ID O75530

AC O75530;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE EMBRYONIC ECTODERM DEVELOPMENT PROTEIN (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 98252946.

RA SEWALT R.G.A.B., VAN DER VLAG J., GUNSTER M.J., HAMER K.M.,

RA DEN BLAAUWEN J.L., SATIJN D.P.E., HENDRIX T., VAN DRIEL R.,

RA OTTE A.P.;

RT \*Characterization of interactions between the mammalian  
 RT polycomb-group proteins Enx1/Ezh2 and EED suggests the existence of  
 RT different mammalian polycomb-group protein complexes.";  
 RL Mol. Cell. Biol. 18:3586-3595(1998).

[2]  
 RN SEQUENCE OF 77-517 FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE; 99026128.  
 RA SCHUMACHER A., LICHTARGE O., SCHWARTZ S., MAGNUSON T.;  
 RT "The murine Polycomb-group gene eed and its human orthologue:  
 RT functional implications of evolutionary conservation.";  
 RL Genomics 54:79-88(1998).  
 DR EMBL; AF070418; AAC23685.1; -;  
 DR EMBL; AF080227; AAC95144.1; -;  
 DR PFAM; PF00400; WD40; 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 517 AA; 57907 MW; 9852BEDE CRC32;

Query Match 68.9%; Score 42; DB 4; Length 517;  
 Best Local Similarity 75.0%; Pred. No. 5.2;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KCKFKFKC 10  
 |||: |||

Db 153 KCKYSFKC 160

RESULT 3  
 P97462  
 ID P97462 PRELIMINARY; PRT; 535 AA.  
 AC P97462;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE (EMBRYONIC ECTODERM DEVELOPMENT  
 GN EED.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA DENISENKO O., BOMSZTYK K.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE OF 95-535 FROM N.A.  
 RC STRAIN=SWISS WEBSTER;  
 RX MEDLINE; 96353973.  
 RA SCHUMACHER A., FAUST C., MAGNUSON T.;  
 RT "Role of leptin in the neuroendocrine response to fasting.";  
 RL Nature 382:250-253(1996).  
 DR EMBL; U97675; AAC53302.1; -;  
 DR EMBL; U78103; AAB38319.1; -;  
 DR MGD; MGI:95286; EED.  
 DR PFAM; PF00400; WD40; 2.  
 SQ SEQUENCE 535 AA; 59997 MW; A2B70C1C CRC32;

Query Match 68.9%; Score 42; DB 11; Length 535;  
 Best Local Similarity 75.0%; Pred. No. 5.3;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KCKFKFKC 10  
 |||: |||

Db 171 KCKYSFKC 178

RESULT 4  
 O59556  
 ID O59556 PRELIMINARY; PRT; 170 AA.  
 AC O59556;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE 170AA LONG HYPOTHETICAL PROTEIN.  
 GN PH1864.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE; 98344137.  
 RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
 RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,  
 RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
 RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
 RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
 RA KIKUCHI H.;  
 RT \*Complete sequence and gene organization of the genome of a  
 RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL; AP000007; BAA30986.1; -;  
 SQ SEQUENCE 170 AA; 19813 MW; F9DECOAL CRC32;

Query Match 67.2%; Score 41; DB 1; Length 170;  
 Best Local Similarity 70.0%; Pred. No. 3.1;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKC 10  
 |||||: |

Db 16 KFKCKFKC 25

RESULT 5  
 Q19807  
 ID Q19807 PRELIMINARY; PRT; 620 AA.  
 AC Q19807;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE SIMILAR TO ZINC FINGER PROTEIN.  
 GN F26A10.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX BENTLEY D.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U28736; AAA68307.1; -;  
 DR PFAM; PF00096; zf-C2H2; 4.

```

DR PROSITE: PS00028; ZINC_FINGER_C2H2; 3.
KW Zinc-finger; Metal-binding; DNA-binding.
SQ SEQUENCE 620 AA; 70218 MW; 586C6C616 CRC32;

Query Match 63.9%; Score 39; DB 5; Length 620;
Best Local Similarity 87.5%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKCKFKFK 9
    |||||
DB 76 FKCKFKFK 83

RESULT 6
ID Q65257 PRELIMINARY; PRT; 66 AA.
AC Q65257;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF L2R.
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage;
OC African swine fever-like viruses.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MALAWI LIL20 /1;
RX MEDLINE: 94014996
RA VYDELINGUM S., BAYLIS S.A., BRISTOW C., SMITH G.L., DIXON L.K.;
RT "duplicated genes within the variable right end of the genome of a
RT pathogenic isolate of African swine fever virus.";
RL J. Gen. Virol. 74:2125-2130(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MALAWI LIL20 /1;
RX MEDLINE: 94292916.
RA DIXON L.K., TWIGG S.R.F., BAYLIS S.A., VYDELINGUM S., BRISTOW C.,
RA HAMMOND J.M., SMITH G.L.;
RT "Nucleotide sequence of a 55 kbp region from the right end of the
RT genome of a pathogenic African swine fever virus isolate (Malawi
RT LIL20/1).";
RL J. Gen. Virol. 7:1655-1684(1994).
DR EMBL: X71982; CAA50854.1;
SQ SEQUENCE 66 AA; 7721 MW; B9756AEB CRC32;

Query Match 63.9%; Score 39; DB 12; Length 66;
Best Local Similarity 66.7%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 KFKCK--FKFK 10
    |||||
DB 17 KFDCKSIFSK 28

RESULT 7
ID Q10956 PRELIMINARY; PRT; 514 AA.
AC Q10956;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-FEB-1997 (TrEMBLrel. 02, Last annotation update)
DE HYPOTHETICAL 59.3 KD PROTEIN B0336.11 IN CHROMOSOME III.
GN B0336.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA TAICH A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
DR EMBL: U32305; AAC46837.1; -.
DR WORMPEP: B0336.11; CE00779.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 449 469 POTENTIAL.
SQ SEQUENCE 514 AA; 59267 MW; E059C05B CRC32;

Query Match 60.7%; Score 37; DB 5; Length 514;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KCKFKFKC 10
    | : || || |
DB 408 KAEKFKFK 415

RESULT 8
ID P91567 PRELIMINARY; PRT; 224 AA.
AC P91567;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE COSMID ZK354.
GN ZK354.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA JOHNSON D., WAMSLEY P., BRADSHAW H.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U88172; AAB42259.1;
SQ SEQUENCE 224 AA; 26534 MW; 6CED87B0 CRC32;

Query Match 60.7%; Score 37; DB 5; Length 224;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KCKFKFKC 10
    || || |
DB 51 KCAFKVKC 58

RESULT 9
ID Q19496 PRELIMINARY; PRT; 548 AA.
AC Q19496;

```

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 63.5 KD PROTEIN.  
 GN F16H11.5  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718;  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX WU X.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; U55376; AAA98005.1; -;  
 DR FRAM; PF00105; zf-C4; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 139 144 POLY-GLN.  
 SQ SEQUENCE 548 AA; 63545 MW; DD7D3BA5 CRC32;

Query Match 60.7%; Score 37; DB 5; Length 548;  
 Best Local Similarity 50.0%; Pred. No. 36;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
 :|||: ||  
 Db 57 RYKCRFRKC 66

RESULT 10

017928 PRELIMINARY; PRT; 338 AA.  
 ID 017928  
 AC 017928  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
 DE NHR-52 PROTEIN.  
 GN NHR-52.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC LLOYD C.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718;  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,

RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 283233; CAB05761.1; -;  
 SQ SEQUENCE 338 AA; 39582 MW; 0F9FB43C CRC32;

Query Match 60.7%; Score 37; DB 5; Length 338;  
 Best Local Similarity 60.0%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KFKCKFKFC 10  
 ||||: ||  
 Db 36 KFTCKYDKC 45

RESULT 11

096867 PRELIMINARY; PRT; 331 AA.  
 ID 096867  
 AC 096867  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE CELL SIGNALING MOLECULE WNT-5 (FRAGMENT).  
 GN SPWNT-5.  
 OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;  
 OC Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98320638.  
 RA FERKOWICZ M.J., STANDER M.C., RAFF R.A.;  
 RT "Phylogenetic relationships and developmental expression of three sea  
 urchin Wnt genes";  
 RL Mol. Biol. Evol. 15:809-819(1998).  
 CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALLING  
 CC MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE REGIONS OF  
 CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SIMILARITY: TO OTHER MEMBERS OF THE WNT FAMILY.  
 DR EMBL; U58982; AAC69436.1; -;  
 DR PROSITE; PS00246; WNT1; 1.  
 KW Developmental protein; Glycoprotein.  
 FT NON\_TER 1  
 SQ SEQUENCE 331 AA; 36989 MW; A94F2700 CRC32;

Query Match 60.7%; Score 37; DB 5; Length 331;  
 Best Local Similarity 75.0%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KCKFKFC 10  
 |||||: |  
 Db 307 KCKFKWCC 314

RESULT 12

09YW73 PRELIMINARY; PRT; 437 AA.  
 ID 09YW73  
 AC 09YW73  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)



```

DE ORF MSV019 HYPOTHETICAL PROTEIN.
GN MSV019.
OS Melanoplus sanguinipes entomopoxvirus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TCUSON;
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus.";
RL J. Virol. 73:533-552(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TCUSON;
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF063866; AAC67851.1; -
SQ SEQUENCE 437 AA; 52581 MW; ADD04712 CRC32;

Query Match 60.7%; Score 37; DB 12; Length 437;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFKCKFKF 8
Db 223 KYCKKFLF 230

RESULT 13
ID O30192 PRELIMINARY; PRT; 363 AA.
AC O30192;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JAN-1998 (TrEMBLrel. 08, Last annotation update)
DE FIRST MANNOSYL TRANSFERASE (WBA2-1).
GN AF0043.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE; 98049343.
RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRIDES N.C.,
RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,
RA COTTON M.D., SPRIGGS T., ARTIACH P., KATNE B.P., SYKES S.M.,
RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA VENTER J.C.;
RT "The complete genome sequence of the hyperthermophilic,
RT sulphate-reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001104; AAB91187.1; -
DR TIGR; AF0043; -
DR PFAM; PF00534; Glycosyl transferase.
KW Hypothetical protein; Transferase.
SQ SEQUENCE 363 AA; 42109 MW; 689B4C2F CRC32;

Query Match 59.0%; Score 36; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KFKCKFKC 10
Db 177 KFKCKFKC 182

```

```

RESULT 14
ID Q09577 PRELIMINARY; PRT; 805 AA.
AC Q09577;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 92.6 KD PROTEIN K02C4.4 IN CHROMOSOME II.
GN K02C4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA LIGHTNING J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS ONE LIM DOMAIN.
CC THE LIM DOMAIN BINDS TO 2 ZINC IONS.
DR EMBL; Z47811; CA87787.1; -
DR WORMPEP; K02C4.4; CE01600.
DR PROSITE; PS00478; LIM_DOMAIN; 1.
DR PFAM; PF00412; LIM; 1.
KW Hypothetical protein; LIM motif; Zinc; Metal-binding.
FT DOMAIN 7 43 LIM.
FT DOMAIN 297 303 POLY-PRO.
SQ SEQUENCE 805 AA; 92562 MW; 17C0D717 CRC32;

Query Match 59.0%; Score 36; DB 5; Length 805;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 KFK--CKFKFKC 10
Db 640 KFKSVCKFKVVC 651

RESULT 15
ID Q05712 PRELIMINARY; PRT; 764 AA.
AC Q05712;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE SUBUNIT OF DNA POLYMERASE II.
GN DPB11.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SKI;
RX MEDLINE; 96102200.
RA ARAKI H., LEWIS S.H., PHONGDARA A., SUGINO A.;
RT "Dpb11, which interacts with DNA polymerase II(epsilon) in
RT Saccharomyces cerevisiae, has a dual role in S-phase progression and
RT at a cell cycle checkpoint.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:11791-11795(1995).
DR EMBL; D42168; BAA07725.1; -
DR PFAM; PF00533; BRCT; 3.
SQ SEQUENCE 764 AA; 87263 MW; BEABDA1F CRC32;

Query Match 57.4%; Score 35; DB 3; Length 764;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FKCKFKFKC 10
Db 595 FKCKIKKPC 603

```

Search completed: September 7, 1999, 20:34:47  
Job time: 19743 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 20:37:17 ; Search time 147.16 Seconds  
(without alignments)  
1.610 Million cell updates/sec

Title: US-09-124-280A-16

Perfect score: 55

Sequence: 1 KLKCKLKLC 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq.36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	55	100.0	10	1 R71787	Peptide neutralisi
2	55	100.0	10	1 W21604	Antibiotic potenti
3	43	78.2	10	1 R71786	Peptide neutralisi
4	43	78.2	10	1 W21603	Antibiotic potenti
5	36	65.5	40	1 R84925	Beta-pleated sheet
6	36	65.5	39	1 W6746	Anticoagulant pept
7	35.5	64.5	11	1 W03066	Polycationic polyp
8	35	63.6	2942	1 W22050	Saccharomyces cere
9	35	63.6	2496	1 W36093	Mutant YLR087c pro
10	34	61.8	551	1 R71376	Human cystathionin
11	34	61.8	77	1 W82858	Antipathogenic pep
12	34	61.8	77	1 W82859	Storage-specific,
13	33	60.0	244	1 P03004	Peptide for treati
14	33	60.0	10	1 R33526	Peptide for treati
15	33	60.0	11	1 R33531	Endotoxin lipid A
16	33	60.0	10	1 R39288	Endotoxin lipid A
17	33	60.0	11	1 R39292	Antibiotic potenti
18	33	60.0	10	1 W21619	Antibiotic potenti
19	33	60.0	11	1 W21623	Human neutrophil g
20	33	60.0	10	1 W62434	Human neutrophil g
21	33	60.0	11	1 W62435	Large polypeptid
22	32	58.2	3457	1 R62504	Aspergillus oryzae
23	32	58.2	604	1 R37163	Polyprotein encode
24	32	58.2	3457	1 W84560	Glutathione reduct
25	32	58.2	479	1 W96320	Antigenic surface
26	31	56.4	970	1 P93305	Sequence of human
27	31	56.4	243	1 P90600	Includes a 6kd hum
28	31	56.4	381	1 R04833	Synthetic hydropho
29	31	56.4	26	1 P80644	Synthetic hydropho
30	31	56.4	78	1 P80645	Human SAPI(53-7
31	31	56.4	26	1 P80589	SAP(Phe). Pulmonar
32	31	56.4	286	1 P82933	Deduced partial se
33	31	56.4	381	1 P80651	Synthetic hydropho
34	31	56.4	52	1 P80646	Sequence of the Ja
35	31	56.4	1664	1 P80264	Human Sp18 deduced
36	31	56.4	382	1 P82982	Human alveolar sur
37	31	56.4	381	1 R06332	Sequence of a huma
38	31	56.4	381	1 P70438	Sequence encoded b
39	31	56.4	61	1 P81076	Sequence of Sp-18
40	31	56.4	381	1 P94756	6kd pulmonary surf
41	31	56.4	381	1 P70664	Gene product of ve
42	31	56.4	381	1 R05093	
43	31	56.4	381	1 R05093	

44 31 56.4 1295 1 R10424 Recombinant vaccin  
45 31 56.4 52 1 R10800 Pulmonary surfacta

ALIGNMENTS

```
RESULT 1
R71787
ID R71787 standard; peptide; 10 AA.
AC R71787;
DT 01-OCT-1995 (first entry)
DE Peptide neutralising toxicity of Lipid A.
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide bond 4..10
PN W09503327-A.
PD 02-FEB-1995.
PF 21-JUL-1994; E02413.
PR 26-JUL-1993; US-097830.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M.
DR WPI; 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising
PT repeating units of a basic amino acid or basic and hydrophobic
PT amino acids
PS Claim 18; Page 21; 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides which
CC include units of formula: (A)n, where A is the cationic amino acid Lys
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or
CC Trp, and p is 2 or greater.
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic
CC activity. Hence they can be used therapeutically to treat septic shock
CC and also in vitro to detoxify vaccines, drug solutions, injectable
CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new
CC peptides.
SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKCKLKLC 10
   |||||
DB 1 KLKCKLKLC 10

RESULT 2
W21604
ID W21604 standard; peptide; 10 AA.
AC W21604;
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #16.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide bond 4..10
PN W09638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M; Varra M.
DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 20; Page 25; 37pp; English.
```

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0095;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 10  
 | | | | | | | | | |  
 Db 1 KLKCKLKLKC 10

## RESULT 3

R71786  
 ID R71786 standard; peptide; 10 AA.

AC R71786; 100.0% (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.

Key Location/Qualifiers  
 FT disulfide\_bond 4..10

PN WO9503327-A.

PD 02-FEB-1995.

PF 21-JUL-1994; E02413.

PR 26-JUL-1993; US-097830.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M;

DR WPI; 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids

PS Claim 17; Page 21; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.

CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.

CC The present sequence is a specifically claimed example of the new  
 CC peptides.

SQ Sequence 10 AA;

Query Match 78.2%; Score 43; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.52;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 10  
 | | | | | | | | | |  
 Db 1 KFKCKFKFKC 10

## RESULT 4

W21603  
 ID W21603 standard; peptide; 10 AA.

AC W21603;

DE 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #15.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.  
 OS Synthetic.

Key Location/Qualifiers  
 FT disulfide\_bond 4..10

PN WO9638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M; Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 19; Page 25; 37pp; English.

CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 78.2%; Score 43; DB 1; Length 10;  
 Best Local Similarity 70.0%; Pred. No. 0.52;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 10  
 | | | | | | | | | |  
 Db 1 KFKCKFKFKC 10

## RESULT 5

R84925

ID R84925 standard; peptide; 40 AA.

AC R84925;

DT 18-MAR-1996 (first entry)

DE Beta-pleated sheet-forming oligopeptide (LKLK)10.

KW Beta-pleated sheet; secondary structure; nucleic acid transfer;

OS cationic; DNA binding peptide; gene therapy; encapsulation.

Key Location/Qualifiers

FT region 1..4

/label= LKLK

/note= "one of 10 repeat units"

PN FR2715847-A1.

PD 11-AUG-1995.

PF 08-FEB-1994; 001381.

PR 08-FEB-1994; FR-001381.

PA (RHON) RHONE POULENC RORER SA.

PI Bazille D, Emile C, Helene C, Spenlehauer G;

DR WPI; 95-276981/37.

PT Complex of nucleic acid and oligopeptide with sec. structure - and  
 PT transfer vectors contg. them, useful for efficient transfer of  
 PT nucleic acid to cells in gene therapy.

PS Claim 6 and Example 1.2; Page 8; 20pp; French.

CC The present peptide is a specific example of a cationic oligopeptide  
 CC corresp. to the formula (b-1-b-1)n, where b is a hydrophobic amino acid,  
 CC 1 is a hydrophilic amino acid and n is at least 4. In this case, where  
 CC b is Leu, 1 is Lys and n = 10, the oligopeptide forms a beta-pleated  
 CC sheet which forms a stable complex with a nucleic acid. The complex  
 CC is suitable for transferring nucleic acid, esp. in gene therapy.  
 SQ Sequence 40 AA;

Query Match 65.5%; Score 36; DB 1; Length 40;  
 Best Local Similarity 88.9%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 9

```

Db      1 11111111
        2 KLKCLKLK 10

RESULT 6
W86746  W86746 standard; peptide; 39 AA.
AC      W86746;
DE      26-MAR-1999 (first entry)
DT      Anticoagulant peptide.
KW      Anticoagulant; blood coagulation inhibitor; disulphide bond; catheter;
KW      blood bag; dialysis membrane; artificial blood vessel.
OS      Synthetic.
FH      Key Location/Qualifiers
FT      Disulfide_bond 14..27
FT      Disulfide_bond 16..25
FN      J11001493-A.
PD      06-JAN-1999.
PF      12-JUN-1997; 172827.
PR      12-JUN-1997; JP-172827.
PA      (KURS ) KURARAY CO LTD.
DR      WPI; 99-125476/11.
PT      New peptide which inhibits blood coagulation - useful in a
PT      pharmaceutical material used as a catheter, blood vessel and blood
PT      dialysis membrane
PS      Disclosure; Page 7; 18pp; Japanese.
CC      New inter- or intra-disulphide bonded peptides are disclosed which have
CC      the formula A'-X'-Cys(1)-Y'-Cys(2)-B' A'-X'-Cys(3)-Y'-Cys(4)-Z'-B'
CC      in which: Cys(1) is disulphide-bonded to Cys(4); Cys(2) is disulphide-
CC      bonded to Cys(3); A' = H or forms a single bond together with B'; B' = OH
CC      or amino or forms a single bond together with A'; A' = H or forms a
CC      single bond together with B'; B' = OH or amino or forms a single bond
CC      together with A'; X and X' = peptide fragments composed of 3 to 13
CC      amino acid residues; Y and Y' = neutral or basic amino acid residues;
CC      and Z and Z' = peptide fragments composed of 2 to 12 amino acid
CC      residues. These peptides inhibit blood coagulation. They can be
CC      immobilised on the blood-contacting surfaces of catheters, blood
CC      circuits, blood bags, blood dialysis membranes, artificial blood
CC      vessels, etc. The present sequence represents a specific example of
CC      the new peptides.
SQ      Sequence 39 AA;

Query Match 65.5%; Score 36; DB 1; Length 39;
Best Local Similarity 88.9%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      1 KLKCLKLK 9
        11111111
Db      1 KLKCLKLK 9

RESULT 7
W03056  W03056 standard; peptide; 11 AA.
AC      W03056;
DE      02-MAR-1997 (first entry)
DT      Polycationic polypeptide component of peptide-oligonucleotide conjugate.
KW      polycationic polypeptide; polyanionic oligonucleotide; antigenic;
KW      antisense therapy.
OS      Synthetic.
FH      Key Location/Qualifiers
FT      modified_site 1
FT      /note= "this residue is connected via the thiol
FT      group and a linking group to the 5' end of
FT      the oligonucleotide T01822"
FT      11
FT      modified_site
FT      /note= "this residue is connected via the thiol
FT      group and a linking group to the 3' end of
FT      the oligonucleotide T01822"
PN      W09524222-A1.
PD      14-SEP-1995.
PF      07-MAR-1995; U02894.

PR      07-MAR-1994; US-207438.
PA      (UYNE-) UNIV NEW JERSEY.
PI      Stein S, Tung C, Wei Z, Zhu T;
DR      WPI; 95-328105/42.
PT      New cyclic conjugate of polycationic polymer and oligo:nucleotide(s)
PT      - covalently bonded at each end by crosslinking agent, useful for
PT      anti:sense and anti:gene therapy, have strong binding to target and
PT      good in-vivo stability
PS      Examples; Page 27; 49pp; English.
CC      Cyclic polycationic polymer-oligonucleotide conjugates are provided
CC      which comprise a polycationic polymer covalently bonded at each end to
CC      the 3' and 5' terminal nucleotides of a polyanionic oligonucleotide via
CC      a crosslinking agent. Preferably the polycationic polymer is a
CC      polypeptide Cys-(Lysleu)2-Lys-(Lysleu)2-Cys (the present sequence) or
CC      Cys-(delta-Orn)10-Cys; the polyanionic oligonucleotide is
CC      5'-CATTTCTTTATT-3' (T01822); and each linking group is of formula
CC      -CH2CONH- where each -CH2 is attached to the thiol group of each Cys
CC      and each NH- is attached to the 5' and 3' terminals respectively of
CC      the oligonucleotide.
CC      The conjugates can be used for antisense and antigene therapy. They
CC      have enhanced stability in-vivo because exonuclease digestion is
CC      minimised when both the 3' and 5' termini are blocked; they have
CC      enhanced bioavailability because the ability of the oligonucleotide
CC      to penetrate through cellular membranes is enhanced when its negative
CC      charges are ion-paired; and they have low toxicity because the
CC      metabolic degradation products of the conjugates are amino acids and
CC      nucleotides.
SQ      Sequence 11 AA;

Query Match 64.5%; Score 35.5; DB 1; Length 11;
Best Local Similarity 90.0%; Pred. No. 7;
Matches 9; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy      1 KLKCLKLK 10
        11111111
Db      3 KLK- KLKCLK 11

RESULT 8
W22050  W22050 standard; Protein; 2942 AA.
AC      W22050;
DT      26-FEB-1998 (first entry)
DE      Saccharomyces cerevisiae protein encoded by novel CSF-1 gene.
KW      Saccharomyces cerevisiae; cold sensitivity of fermentability; CSF-1;
KW      low temperature sensitivity; yeast; bread; ethanol; fermentation.
OS      Saccharomyces cerevisiae.
PN      W09724442-A1.
PD      10-JUL-1997.
PF      27-DEC-1996; J03862.
PR      28-DEC-1995; JP-343700.
PA      (KYOW ) KYOWA HAKKO KOGYO KK.
PI      Kawasaki H, Kikuchi Y, Ouchi K, Tokai M;
DR      WPI; 97-363678/33.
DR      N-PSDB; T74989.
PT      Protein for complementing low temperature sensitivity in
PT      fermentation - can be inactivated in yeast and used for producing
PT      bread and ethanol
PS      Claim 3; Pages 21-40; 48pp; Japanese.
CC      This is a protein capable of complementing a mutation showing low
CC      temperature sensitivity in fermentation. This is encoded by a novel
CC      genomic DNA CSF-1 (cold sensitivity of fermentability). The Saccharomyces
CC      cerevisiae YHK 1243 has an inactivated gene encoding this new protein or
CC      has the novel DNA sequence contained in its chromosome. This protein can
CC      be used in dough to produce bread, and to produce ethanol.
SQ      Sequence 2942 AA;

Query Match 63.6%; Score 35; DB 1; Length 2942;
Best Local Similarity 50.0%; Pred. No. 9.6e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 KLKCKLKLKC 10
   || || :||
Db 126 KLPCCKISVEC 135

RESULT 9
W36093
ID W36093 standard; Protein; 2496 AA.
AC W36093;
DE 25-MAR-1998 (first entry)
DE Mutant YLR087c protein from cold sensitive yeast strain.
KW Mutant; cold sensitive; yeast; complementation; breadmaking; dough;
KW sugar; leavening agent; brewing; winemaking; truncation.
OS Saccharomyces cerevisiae.
FH Key Location/Qualifiers
FT Misc_difference 527 /note= "encoded by GCR"
FT Misc_difference 618 /note= "encoded by GAY"
FT Misc_difference 640 /note= "encoded by TCY"
FT Misc_difference 1103 /note= "encoded by AGR"
FT Misc_difference 1582 /note= "encoded by GAR"
PN W09728693-A1.
PD 14-AUG-1997.
PF 07-FEB-1997; F00254.
PR (08-FEB-1996; FR-001562.
PA (LESA ) LESAFFRE & CIE.
PI Colavizza D, Loiez A, Wadoux I;
DR WPI: 97-414988/38.
DR N-PSDB; T94548.
PT New strains of bread-making yeast with low fermentative activity at
PT low temperature - allows production of doughs that can be stored
PT cold for many hours before final baking
PS Example 4; Page 59-64; 76pp; French.
CC This is the amino acid sequence of a mutant protein (designated YLR087c)
CC isolated from a cold sensitive strain of Saccharomyces cerevisiae. The
CC gene sequence was isolated from the clone YCP50-10.39 which was able to
CC complement the cold sensitive strain HLI3.2.30. This protein is a
CC truncated mutation as compared to the wild type sequence. The wild type
CC gene contains an open reading frame of 8874 bases encoding a protein of
CC 2958 amino acids. The mutant gene has an open reading frame of 7488
CC bases encoding a protein of 2496 amino acids. Yeast strains containing
CC the YLR087c gene are used, fresh or dried, in modified breadmaking
CC processes, i.e. in processes where the delay between mixing the dough
CC and baking exceeds 6 hr. Particularly they are used to make French-style
CC bread (no added sugar) or breads with sugar content below 5%. The strains
CC can also be used to make mixed yeast/bacteria acidic leavening agents.
CC The strains can also be used in brewing and winemaking as a
CC cold-sensitive phenotype.
SQ Sequence 2496 AA;

Query Match 63.6%; Score 35; DB 1; Length 2496;
Best Local Similarity 50.0%; Pred. No. 8.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 10
   || || :||
Db 126 KLPCCKISVEC 135

RESULT 10
R71376
ID R71376 standard; Protein; 551 AA.
AC R71376;
DE 21-NOV-1995 (first entry)
DE Human cystathionine beta-synthase.
KW Cystathionine; beta synthase; human; homocystinuria.
OS Homo sapiens.
PN W09507714-A.

```

```

PD 23-MAR-1995.
PF 12-SEP-1994; U10203.
PR 13-SEP-1993; US-120960.
PA (COLS ) UNIV COLORADO.
PI Kraus JP;
DR WPI: 95-131185/17.
DR N-PSDB; Q87430.
PT Purified DNA encoding human cystathionine beta-synthase - useful
PT for producing human cystathionine beta-synthase, used for
PT treating homocystinuria.
PS Claim 7; Fig 1a-c; 45pp; English.
CC The amino acid sequence shown in R71376 is the human cystathionine
CC beta-synthase (CBS). Human CBS can be used to treat patients with
CC homocystinuria and the cDNA sequence, Q87430, from which it is
CC derived is also useful for screening CBS deficient patients for
CC mutations in the CBS gene.
SQ Sequence 551 AA;

Query Match 61.8%; Score 34; DB 1; Length 551;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLKC 10
   |||:| ||
Db 101 LKCELLAKC 109

RESULT 11
W82858
ID W82858 standard; peptide; 77 AA.
AC W82858;
DE 19-MAY-1999 (first entry)
DE Antipathogenic peptide.
KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
KW cancer; infection; disinfectant; contact lens wetting solution;
KW preservative; pesticide; fungicide; bactericide.
OS Synthetic.
PN W09837090-A1.
PD 27-AUG-1998.
PF 19-FEB-1998; IL0081.
PR 20-FEB-1997; WO-IL0066.
PA (YEDA ) YEDA RES & DEV CO LTD.
PI Oren Z, Shai Y;
DR WPI: 98-594464/50.
PT New non-haemolytic cytolytic agent useful in treating cancer or
PT infections - is a peptide comprising a moiety which disrupts the
PT continuity of an alpha-helical structure
PS Claim 17; Page 106; 126pp; English.
CC The present peptide is used to produce the agents of the invention. The
CC specification describes a non-haemolytic, cytolytic agent, which is a
CC peptide, a complex of bundled peptides, a mixture of peptides or a random
CC peptide copolymer. The agent has a selective cytolytic activity on
CC pathogenic cells. The agent is selected from a cyclic derivative of a
CC peptide which has a net positive charge greater than 1, comprises L-amino
CC acid residues and/or D-amino acid residues and comprises an alpha-helix
CC breaker moiety, or a peptide (or cyclic derivative of this) which
CC (comprises L-amino acid residues and D-amino acid residues, has a net
CC positive charge greater than 1 and has an amino acid sequence such that
CC a corresponding amino acid sequence comprising only L-amino acid residues
CC is not found in nature. The cytolytic agents may be used for treatment of
CC cancer or for treatment of several diseases caused by pathogens,
CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
CC They may be used in both human and veterinary medicine. They may also be
CC used as disinfectants for destruction of microorganisms, i.e. in the
CC solutions for wetting contact lenses, as preservatives, e.g., in the
CC cosmetic and food industries, as pesticides (e.g. fungicides or
CC bactericides) or for preservation of agricultural products.
SQ Sequence 77 AA;

Query Match 61.8%; Score 34; DB 1; Length 77;
Best Local Similarity 87.5%; Pred. No. 60;

```

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKCKLKLK 9  
| | | | |  
Db 11 LKCKLKLK 18

## RESULT 12

W82859  
ID W82859 standard; peptide: 77 AA.

AC W82859;

DT 19-MAY-1999 (first entry)

DE Antipathogenic peptide.

KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;

KW cancer; infection; disinfectant; contact lens wetting solution;

KW preservative; pesticide; fungicide; bactericide.

OS Synthetic.

PN W09837090-A1.

PD 27-AUG-1998.

PF 19-FEB-1998; IL0081.

PR 20-FEB-1997; WO-IL0066.

PA (YEDA ) YEDA RES & DEV CO LTD.

PI Oren Z, Shai Y;

DR WPI: 98-594464/50.

PT New non-haemolytic cytolytic agent useful in treating cancer or

PT infections - is a peptide comprising a moiety which disrupts the

PT continuity of an alpha-helical structure

PS Claim 17; Page 107; 126pp; English.

CC The present peptide is used to produce the agents of the invention. The

CC specification describes a non-haemolytic, cytolytic agent, which is a

CC peptide, a complex of bundled peptides, a mixture of peptides or a random

CC peptide copolymer. The agent has a selective cytolytic activity on

CC pathogenic cells. The agent is selected from a cyclic derivative of a

CC peptide which has a net positive charge greater than 1, comprises L-amino

CC acid residues and/or D-amino acid residues and comprises an alpha-helix

CC breaker moiety, or a peptide (or cyclic derivative of this) which

CC comprises L-amino acid residues and D-amino acid residues, has a net

CC positive charge greater than 1 and has an amino acid sequence such that

CC a corresponding amino acid sequence comprising only L-amino acid residues

CC is not found in nature. The cytolytic agents may be used for treatment of

CC cancer or for treatment of several diseases caused by pathogens,

CC including bacterial, fungal, viral, mycoplasma and protozoan infections.

CC They may be used in both human and veterinary medicine. They may also be

CC used as disinfectants for destruction of microorganisms, i.e. in

CC solutions for wetting contact lenses, as preservatives, e.g., in the

CC cosmetic and food industries, as pesticides (e.g. fungicides or

CC bactericides) or for preservation of agricultural products.

CC Sequence 77 AA;

SQ

Query Match 61.8%; Score 34; DB 1; Length 77;

Best Local Similarity 87.5%; Pred. No. 60;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKCKLKLK 9

| | | | |

Db 11 LKCKLKLK 18

## RESULT 13

P50304

ID P50304 standard; Protein: 244 AA.

AC P50304;

DT 22-OCT-1991 (first entry)

DE Storage-specific, late schizont merozoite malaria antigen.

KW Malaria; vaccine.

OS Plasmodium spp.

PN W08503725-A.

PD 29-AUG-1985.

PF 20-FEB-1985; G00072.

PR 20-FEB-1984; GB-004378.

PA (BIOU ) BIOGEN NV.

PI Mach B, Perrin L, McGarvey M, Cheung A, Shaw A;

DR WPI: 85-223371/36.

DR N-PSDB; N50355.

PT Prodn. of antigens of Plasmodium species, esp. of falciparum - by

PT recombinant DNA methods giving polypeptide(s) for protecting

PT against malaria or for diagnosis.

PS Disclosure; Fig 6; 49pp; English.

CC The sequence encodes a Plasmodium falciparum, Plasmodium vivax,

CC Plasmodium malariae and Plasmodium ovale antigen which may be used

CC in the diagnosis of malaria and as a vaccine against malaria.

SQ Sequence 244 AA;

Query Match

Best Local Similarity 60.0%; Score 33; DB 1; Length 244;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLKCKLKLK 10

| | | | |

Db 33 KANCKCKNKC 42

## RESULT 14

R33526

ID R33526 standard; peptide; 10 AA.

AC R33526;

DE 07-JUL-1993 (first entry)

DT Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;

KW cytokine release control; treatment; pertussis; bacterial meningitis;

KW HIV related infections; polymyxin B.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide\_bond 4. 10

PN ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORR M.

PI PORR M;

DR WPI: 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises

PT specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated

PS Claim 4; Page 32; 39pp; English.

CC This is a specific example of a generic peptide of formula

CC R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock,

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

SQ Sequence 10 AA;

Query Match

Best Local Similarity 60.0%; Score 33; DB 1; Length 10;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLKCKLKLK 10

| | | | |

Db 1 KTKCKFLKNC 10

## RESULT 15

R33531  
 ID R33531 standard; peptide; 11 AA.  
 AC R33531;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 5..11  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PE 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M.  
 DR WPI; 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 8; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 60.0%; Score 33; DB 1; Length 11;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLKCKLKLKC 10  
 | | | | |  
 Db 2 KTKCKFLKLC 11

Search completed: September 7, 1999, 20:37:18  
 Job time: 18467 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:38:37 ; Search time 80.79 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-16

Perfect score: 55

Sequence: 1 KLKCKLKLC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	1	US-08-097-830E-16
2	55	100.0	10	2	US-08-456-112B-16
3	43	78.2	10	1	US-08-097-830E-15
4	43	78.2	10	2	US-08-456-112B-15
5	40	72.7	11	2	US-08-456-112B-13
6	37	67.3	1377	2	US-08-308-818-4
7	36	65.5	40	2	US-08-687-551-7
8	36	65.5	391	2	US-08-644-034A-1
9	34	61.8	551	1	US-08-120-960-2
10	33	60.0	10	1	US-08-049-871-2
11	33	60.0	11	1	US-08-049-871-6
12	33	60.0	10	1	US-07-819-893-2
13	33	60.0	11	1	US-07-819-893-6
14	33	60.0	10	1	US-08-280-397-2
15	33	60.0	11	1	US-08-280-397-6
16	33	60.0	10	2	US-08-218-026-49
17	33	60.0	11	2	US-08-218-026-50
18	33	60.0	10	2	US-08-653-632-49
19	33	60.0	11	2	US-08-653-632-50
20	33	60.0	10	2	US-08-456-112B-31
21	33	60.0	11	2	US-08-456-112B-35
22	32	58.2	3457	2	US-08-416-603-4
23	31	56.4	26	1	US-08-296-898-1
24	31	56.4	463	1	US-08-142-439A-2
25	31	56.4	621	1	US-08-208-887A-49
26	31	56.4	26	2	US-08-435-019-2
27	31	56.4	13	2	US-08-435-019-10
28	31	56.4	16	2	US-08-435-019-17
29	31	56.4	536	2	US-08-890-094-2
30	31	56.4	548	2	US-08-890-094-18
31	31	56.4	463	2	US-08-869-477-2
32	31	56.4	27	3	PCT-US95-09338-70
33	31	56.4	27	3	PCT-US95-09339-70
34	30.5	55.5	881	1	US-08-333-901-1
35	30.5	55.5	881	1	US-08-456-582-1
36	30	54.5	589	1	US-07-668-648-2
37	30	54.5	1253	1	US-08-252-966B-12
38	30	54.5	1261	1	US-08-252-966B-18
39	30	54.5	132	1	US-08-304-051-21

Sequence 4, Appl  
Sequence 14, Appl  
Sequence 4, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-08-097-830E-16  
; Sequence 16, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; FEATURE:  
; OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-16

Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLKCKLKLC 10  
Db 1 KLKCKLKLC 10

RESULT 2

US-08-456-112B-16  
; Sequence 16, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION NUMBER:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;;  
US-08-456-112B-16

Query Match 100.0%; Score 55; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLCKKLKLC 10  
| | | | | | | | | |  
DB 1 KLCKKLKLC 10

RESULT 3  
US-08-097-830E-15  
; Sequence 15, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.

;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; FEATURE:  
;; OTHER INFORMATION: sulfide bond between Cys and  
;; OTHER INFORMATION: Cys  
US-08-097-830E-15

Query Match 78.2%; Score 43; DB 1; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.24;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLCKKLKLC 10  
| | | | | | | | | |  
DB 1 KFKCKFKKC 10

RESULT 4  
US-08-456-112B-15  
; Sequence 15, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-15

Query Match 78.2%; Score 43; DB 2; Length 10;  
Best Local Similarity 70.0%; Pred. No. 0.24;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLCKKLKLC 10  
| | | | | | | | | |

APPLICATION NUMBER: WO PCT/FR95/00098  
FILING DATE: 27-JAN-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: 38,619  
REFERENCE/DOCKET NUMBER: ST94007-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610)454-3839  
TELEFAX: (610)454-3808  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-687-551-7

Query Match 65.5%; Score 36; DB 2; Length 40;  
Best Local Similarity 88.9%; Pred. No. 8.5;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKCKLKLK 9  
DB 2 LKCKLKLK 10

RESULT 8  
US-08-644-034A-1  
Sequence 1, Application US/08644034A  
Patent No. 5882862  
GENERAL INFORMATION:  
APPLICANT: Howard B. Lieberman and Kevin M. Hopkins  
TITLE OF INVENTION: A HUMAN RADIORESISTANCE/CELL CYCLE  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham, LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/644,034A  
FILING DATE: May 9, 1996  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: John P. White  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/50307  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0526  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Amino Acid  
US-08-644-034A-1

Query Match 65.5%; Score 36; DB 2; Length 391;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKCKLKLK 10

DB 71 LRCKILMK 79  
I:::|

RESULT 9  
US-08-120-960-2  
Sequence 2, Application US/08120960  
Patent No. 5523225  
GENERAL INFORMATION:  
APPLICANT: KRAUS, JAN P  
TITLE OF INVENTION: DNA SEQUENCE ENCODING HUMAN  
CYSTATHIONINE B-SYNTHASE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DILWORTH & BARRESE  
STREET: 4350 LA JOLLA VILLAGE DRIVE, SUITE 300  
CITY: SAN DIEGO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/120,960  
FILING DATE: 12-SEP-1993  
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: PEPPER PH.D., FREDERICK W.  
REGISTRATION NUMBER: 31,286  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-546-4410  
TELEFAX: 619-453-2839  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 551 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-120-960-2

Query Match 61.8%; Score 34; DB 1; Length 551;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLK 10  
DB 101 LKCELLAK 109  
I:::|

RESULT 10  
US-08-049-871-2  
Sequence 2, Application US/08049871  
Patent No. 5358933  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
of Bacterial Endotoxins and for the  
Prevention and Treatment of Septic  
Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-2

Query Match 60.0%; Score 33; DB 1; Length 10;  
Best Local Similarity 60.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLKCKLKLKC 10  
| | | | |  
Db 1 KTKCKFLKKC 10

RESULT 11  
US-08-049-871-6  
; Sequence 6, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porzio, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-6

Query Match 60.0%; Score 33; DB 1; Length 11;  
Best Local Similarity 60.0%; Pred. No. 7.6;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLKCKLKLKC 10  
| | | | |  
Db 2 KTKCKFLKKC 11

RESULT 12  
US-07-819-893-2  
; Sequence 2, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porzio, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: circular  
US-07-819-893-2

Query Match 60.0%; Score 33; DB 1; Length 10;  
Best Local Similarity 60.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLKCKLKLKC 10  
| | | | |  
Db 1 KTKCKFLKKC 10

RESULT 13  
US-07-819-893-6  
; Sequence 6, Application US/07819893  
; Patent No. 5371186

GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/819,893  
FILING DATE: 19920115  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: circular  
US-07-819-893-6

Query Match 60.0%; Score 33; DB 1; Length 11;  
Best Local Similarity 60.0%; Pred. No. 7.6;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLKCKLKLKC 10  
Db 2 KTKCKFLKCC 11

RESULT 14  
US-08-280-397-2  
Sequence 2, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-280-397-2

Query Match 60.0%; Score 33; DB 1; Length 10;  
Best Local Similarity 60.0%; Pred. No. 7;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLKCKLKLKC 10  
Db 1 KTKCKFLKCC 10

RESULT 15  
US-08-280-397-6  
Sequence 6, Application US/08280397  
Patent No. 5589459  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/280,397  
FILING DATE: 07/26/94  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/819,893  
FILING DATE: 01/16/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-002A  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular

US-08-280-397-6

Query Match 60.0%; Score 33; DB 1; Length 11;  
Best Local Similarity 60.0%; Pred. NO. 7.6;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKCKLKLKC 10  
| | | | |  
Db 2 KTKCKFLKKC 11

Search completed: September 7, 1999, 22:38:37  
Job time: 7921 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:06:29 ; Search time 116.8 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-16  
Perfect score: 55  
Sequence: 1 KLKCKLKLC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	74.5	246	A64326	hypothetical prote
2	39	70.9	534	C70842	probable pmmb prot
3	38	69.1	534	S73037	probable phosphom
4	37	67.3	2068	A47371	transcription init
5	37	67.3	1490	S32373	DNA-binding protei
6	35	63.6	1969	S02771	myosin heavy chain
7	35	63.6	2958	S64921	probable membrane
8	34	61.8	552	A55760	cystathionine beta
9	34	61.8	561	A42790	cystathionine beta
10	34	61.8	478	C42790	cystathionine beta
11	34	61.8	546	B42790	cystathionine beta
12	34	61.8	547	JX0145	hemoprotein H-450
13	34	61.8	951	T01377	hypothetical prote
14	33	60.0	26926	I38344	titin, cardiac mus
15	33	60.0	500	ISECAB	L-arabinose isomer
16	33	60.0	500	ISEBAB	L-arabinose isomer
17	33	60.0	39	C69577	phosphatase (RapF)
18	33	60.0	501	A43302	probable finger pr
19	33	60.0	215	C73390	transfer protein T
20	33	60.0	209	F64416	polyferredoxin - M
21	33	60.0	776	S59790	hypothetical prote
22	33	60.0	2539	B71619	hypothetical prote
23	33	60.0	661	I52603	MPS1 protein - mou
24	33	60.0	170	C71199	hypothetical prote
25	32	58.2	479	RDHUU	glutathione reduct
26	32	58.2	2330	RRIMV	genome polyprotein
27	32	58.2	369	ESBYPC	3',5'-cyclic-nucle
28	32	58.2	42	MXRSMV	myotoxin a 6 - pra
29	32	58.2	43	CXRSCH	toxic peptide C -
30	32	58.2	466	S39494	glutathione reduct
31	32	58.2	456	A64052	glutathione reduct
32	32	58.2	2331	A44054	genome polyprotein
33	32	58.2	551	S72485	peptidylprolyl iso
34	32	58.2	559	S55383	peptidylprolyl iso
35	32	58.2	43	B29089	myotoxin II - midg
36	32	58.2	43	A29089	myotoxin I - midg
37	32	58.2	45	S12509	myotoxin - western
38	32	58.2	65	JCS324	myotoxin a precurs
39	32	58.2	448	H69032	polyferredoxin 2 -

40	32	58.2	610	2	S58885	derailed receptor
41	32	58.2	136	2	E69480	conserved hypothet
42	32	58.2	1424	2	S11480	hypothetical prote
43	32	58.2	442	2	C03172	helicase homolog 0
44	32	58.2	205	2	C71691	ribosomal protein
45	32	58.2	639	2	H70661	probable dnaG prot

ALIGNMENTS

RESULT 1	
A64326	hypothetical protein MJ0208 - Methanococcus jannaschii
C:Species:	Methanococcus jannaschii
C>Date:	13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 12-Feb-1999
C:Accession:	A64326
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak	
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.	
Science 273, 1058-1073, 1996	
A:Authors:	Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A>Title:	Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number:	A64300; MUID:96337999
A:Accession:	A64326
A>Status:	preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type:	DNA
A:Residues:	1-246 <BUL>
A:Cross-references:	GB:U67476; GB:L77117; NID:gl590942; PID:gl498983; TIGR:MJ0208
C:Genetics:	
A:Map position:	REV200166-199426
C:Superfamily:	hypothetical protein MJ0208; ferredoxin 2[4Fe-4S] homology
F:152-206/Domain:	ferredoxin 2[4Fe-4S] homology <FER>

Query Match	74.5%	Score 41;	DB 2;	Length 246;
Best Local Similarity	80.0%;	Pred. No. 4.3;		
Matches 8;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

QY 1	KLKCKLKLC 10
DB 156	KKKCKLCLKC 165

RESULT 2	
C70842	probable pmmb protein - Mycobacterium tuberculosis (strain H37RV)
C:Species:	Mycobacterium tuberculosis
C>Date:	17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
C:Accession:	C70842
R:Coie, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon	
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,	
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.	
Nature 393, 537-544, 1998	
A:Authors:	Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title:	Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number:	A70500; MUID:98295987
A:Accession:	C70842
A>Status:	preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type:	DNA
A:Residues:	1-534 <COL>
A:Cross-references:	GB:AL021841; GB:AL123456; NID:g3261517; PID:e1251139; PID:g289421
A:Experimental source:	strain H37RV
C:Genetics:	
A:Gene:	pmmb

Query Match	70.9%	Score 39;	DB 2;	Length 534;
Best Local Similarity	60.0%;	Pred. No. 17;		
Matches 6;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1	KLKCKLKLC 10

Db 497 KLCYLEIRC 506

RESULT 3

S73037  
probable phosphomannomutase (EC 5.4.2.8) - Mycobacterium leprae  
N:Alternate names: L308\_Fl\_13 protein  
C:Species: Mycobacterium leprae  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Sep-1997  
C:Accession: S73037  
R:Smith, D.R.; Robison, K.  
submitted to the EMBL Data Library, November 1993  
A:Description: Mycobacterium leprae cosmid L308.  
A:Reference number: S72580  
A:Accession: S73037  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-534 <SMI>  
A:Cross-references: EMBL:U00022; NID:g467164; PID:g467178  
C:Genetics:  
A:Start codon: GTG  
C:Keywords: intramolecular transferase; isomerase

Query Match 69.1%; Score 38; DB 2; Length 534;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLCKKLLKLC 10

|||||I:::|

Db 498 KLCYLEVRC 507

RESULT 4

A47371  
transcription initiation factor IID 230K chain - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
C:Accession: A47371  
R:Kokubo, T.; Gong, D.W.; Yamashita, S.; Horikoshi, M.; Roeder, R.G.; Nakatani, Y.  
Genes Dev. 7, 1033-1046, 1993  
A:Title: Drosophila 230-kD TFIID subunit, a functional homolog of the human cell cycle  
A:Reference number: A47371; MUID:93279463  
A:Accession: A47371  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-2068 <KOK>  
A:Cross-references: GB:S61883; NID:g385550; PID:g385551  
A:Note: sequence inconsistent with nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:133002, NCBIP:133003)  
C:Genetics:  
A:Gene: FlyBase:Taf250  
A:Cross-references: FlyBase:FBgn0010355  
C:Superfamily: unassigned bromodomain proteins; bromodomain homology  
F:1498-1553/Domain: bromodomain homology <BRO1>  
F:1620-1675/Domain: bromodomain homology <BRO>

Query Match 67.3%; Score 37; DB 2; Length 2068;  
Best Local Similarity 80.0%; Pred. No. 11e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLCKKLLKLC 10

|||||I:::|

Db 1342 KLKPDLLKLC 1351

RESULT 5

S32373  
DNA-binding protein TAF-II 250K - fruit fly (Drosophila sp.) (fragment)  
C:Species: Drosophila sp.  
C:Date: 20-May-1994 #sequence\_revision 01-Sep-1995 #text\_change 12-Sep-1997  
C:Accession: S32373

R.Weinzierl, R.O.J.; Dynlacht, B.D.; Tjian, R.  
Nature 362, 511-517, 1993  
A:Title: Largest subunit of Drosophila transcription factor IID directs assembly of a  
A:Reference number: S32373  
A:Accession: S32373  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1490 <WEI>  
C:Genetics:  
A:Gene: FlyBase:Taf250  
A:Cross-references: FlyBase:FBgn0010355  
C:Superfamily: unassigned bromodomain proteins; bromodomain homology  
F:838-901/Domain: bromodomain homology <BRO1>  
F:969-1024/Domain: bromodomain homology <BRO>

Query Match 67.3%; Score 37; DB 2; Length 1490;  
Best Local Similarity 80.0%; Pred. No. 82;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLCKKLLKLC 10

|||||I:::|

Db 682 KLKPDLLKLC 691

RESULT 6

S02771  
myosin heavy chain A - Caenorhabditis elegans  
N:Contains: myosin ATPase (EC 3.6.1.32)  
C:Species: Caenorhabditis elegans  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 13-Feb-1998  
C:Accession: S02771  
R:Diibb, N.J.; Maruyama, I.N.; Krause, M.; Karn, J.  
J. Mol. Biol. 205, 603-613, 1989  
A:Title: Sequence analysis of the complete Caenorhabditis elegans myosin heavy chain  
A:Reference number: S02771; MUID:89178677  
A:Accession: S02771  
A:Molecule type: DNA  
A:Status: nucleic acid sequence not shown  
A:Residues: 1-1969 <DTB>  
A:Cross-references: EMBL:X08067; NID:g6798; PID:g6799  
C:Genetics:  
A:Gene: myo-3  
A:Introns: 46/1; 116/3; 169/1; 269/1; 445/2; 1898/3  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle  
F:89-779/Domain: myosin motor domain homology <MMOT>  
F:179-186/Region: nucleotide-binding motif A (P-loop)  
F:667-689/Region: actin binding #status predicted  
F:770-784/Region: actin binding #status predicted  
F:852-1969/Domain: coiled coil #status predicted <COI>  
F:852-1166/Region: S2  
F:1167-1969/Region: light meromyosin  
F:130/Modified site: N6,N6-trimethyllysine (Lys) #status predicted  
F:185/Binding site: ATP (Lys) #status predicted  
F:707.717/Active site: Cys #status predicted

Query Match 63.6%; Score 35; DB 1; Length 1969;  
Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLCKKLLKLC 8

|||||I:::|

Db 1883 KLQCKLKI 1890

RESULT 7

S64921  
probable membrane protein YLR087c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein L2506  
C:Species: Saccharomyces cerevisiae  
C:Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 21-Nov-1997

C:Accession: S64921; S64919  
R:Benes, V.; Rechmann, S.; Nentwich, U.; Schwager, C.; Ansoerge, W.; Voss, H.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64920  
A:Accession: S64921  
A:Molecule type: DNA  
A:Residues: 1-2958 <BEN>  
A:Cross-references: EMBL:273259; NID:gl360456; PID:e245797; PID:gl360457; MIPS:YLR087c  
A:Experimental source: strain S288C  
R:Pohl, T.M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64899  
A:Accession: S64919  
A:Molecule type: DNA  
A:Residues: 2353-2958 <POH>  
A:Cross-references: EMBL:273259; MIPS:YLR087c  
A:Experimental source: strain S288C  
C:Genetics:  
A:Map position: 12R  
C:Keywords: transmembrane protein  
F:21-37/Domain: transmembrane #status predicted <TM1>  
F:1219-1235/Domain: transmembrane #status predicted <TM2>  
F:1928-1944/Domain: transmembrane #status predicted <TM3>  
F:2654-2670/Domain: transmembrane #status predicted <TM4>

Query Match 63.6%; Score 35; DB 2; Length 2958;  
Best Local Similarity 50.0%; Pred. No. 3e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KUKCKLKLC 10  
|||:|:|  
Db 126 KLPCKISVEC 135

RESULT 8  
A55760  
cystathionine beta-synthase (EC 4.2.1.22) - human  
N:Alternate names: beta-thionase; methylcysteine synthase; serine sulphydrase  
C:Species: Homo sapiens (man)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 16-Dec-1998  
C:Accession: A55760; S49459; I54342  
R:Kruger, W.D.; Cox, D.R.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6614-6618, 1994  
A:Title: A yeast system for expression of human cystathionine beta-synthase: structural  
A:Reference number: A55760; MUID:94294429  
A:Accession: A55760  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-552 <KRU>  
R:Chasse, J.F.; Paris, D.; Paly, E.; Kamoun, P.; London, J.  
submitted to the EMBL Data Library, October 1994  
A:Description: Characterization of human cystathionine beta-synthase (CBS) cDNAs: evidence  
A:Reference number: S49459  
A:Accession: S49459  
A:Molecule type: mRNA  
A:Residues: 1-360, 'A', 362-364, 366-418, 'G', 420-492, 'L', 494-498, 'R', 500-543, 'V', 545-552 <C>  
A:Cross-references: EMBL:X82166; NID:g558581; PID:g558582  
R:Kraus, J.P.; Le, K.; Swaroop, M.; Ohura, T.; Tahara, T.; Rosenberg, L.E.; Roper, M.D.;  
Hum. Mol. Genet. 2, 1633-1638, 1993  
A:Title: Human cystathionine beta-synthase cDNA: sequence, alternative splicing and expression  
A:Reference number: I54342; MUID:94093551  
A:Accession: I54342  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-360, 'A', 362-364, 366-418, 'G', 420-492, 'L', 494-498, 'R', 500-543, 'V', 545-552 <R>  
A:Cross-references: GB:L19501; NID:g388715; PID:g388716  
C:Genetics:  
A:Gene: GDB:CBS  
A:Cross-references: GDB:119754; OMIM:236200  
A:Map position: 21q22.3-21q22.3  
C:Superfamily: cystathionine beta-synthase: CBS homology

C:Keywords: carbon-oxygen lyase; hydro-lyase  
F:422-470/Domain: CBS homology <CBS2>

Query Match 61.8%; Score 34; DB 2; Length 552;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLC 10  
|||:|:|  
Db 101 LKCELLAKC 109

RESULT 9  
A42790  
cystathionine beta-synthase (EC 4.2.1.22), splice form I - rat  
N:Alternate names: beta-thionase; methylcysteine synthase; serine sulphydrase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 15-Jan-1999  
C:Accession: A42790  
R:Swaroop, M.; Bradley, K.; Ohura, T.; Tahara, T.; Roper, M.D.; Rosenberg, L.E.; Krau  
J. Biol. Chem. 267, 11455-11461, 1992  
A:Title: Rat cystathionine beta-synthase. Gene organization and alternative splicing.  
A:Reference number: A42790; MUID:92283859  
A:Accession: A42790  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-561 <SWA>  
A:Cross-references: GB:M88344; NID:gl364273; PID:g206597  
A:Experimental source: strain Sprague-Dawley; liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:104763)  
C:Superfamily: cystathionine beta-synthase: CBS homology  
C:Keywords: alternative splicing; carbon-oxygen lyase; hydro-lyase  
F:417-465/Domain: CBS homology <CBS2>

Query Match 61.8%; Score 34; DB 2; Length 561;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLC 10  
|||:|:|  
Db 98 LKCELLAKC 106

RESULT 10  
C42790  
cystathionine beta-synthase (EC 4.2.1.22), splice form IV - rat  
N:Alternate names: beta-thionase; methylcysteine synthase; serine sulphydrase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 16-Dec-1998  
C:Accession: C42790  
R:Swaroop, M.; Bradley, K.; Ohura, T.; Tahara, T.; Roper, M.D.; Rosenberg, L.E.; Krau  
J. Biol. Chem. 267, 11455-11461, 1992  
A:Title: Rat cystathionine beta-synthase. Gene organization and alternative splicing.  
A:Reference number: A42790; MUID:92283859  
A:Accession: C42790  
A:Status: preliminary  
A:Molecule type: nucleic acid; protein  
A:Residues: 1-478 <SWA>  
A:Experimental source: liver  
C:Superfamily: cystathionine beta-synthase: CBS homology  
C:Keywords: alternative splicing; carbon-oxygen lyase; cysteine biosynthesis; hydro-1

Query Match 61.8%; Score 34; DB 2; Length 478;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLC 10  
|||:|:|  
Db 98 LKCELLAKC 106

## RESULT 11

B42790  
cystathionine beta-synthase (EC 4.2.1.22), splice form III - rat  
N:Alternate names: beta-thionase; methylcysteine synthase; serine sulphydrase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 15-Jan-1999  
C:Accession: B42790  
R:Swarcop, M.; Bradley, K.; Ohura, T.; Tahara, T.; Roper, M.D.; Rosenberg, L.E.; Kraus, J. Biol. Chem. 267, 11455-11461, 1992  
A:Title: Rat cystathionine beta-synthase. Gene organization and alternative splicing.  
A:Reference number: B42790; MUID:92283859  
A:Accession: B42790  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-546 <SWA>  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:104768)  
C:Superfamily: cystathionine beta-synthase; CBS homology  
C:Keywords: alternative splicing; carbon-oxygen lyase; cysteine biosynthesis; hydro-lyase  
F:417-465/Domain: CBS homology <CBS2>

Query Match 61.8%; Score 34; DB 2; Length 546;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

## QY 2 LKCKLKLC 10

|||||  
|||:| ||

## Db 98 LKCELLAKC 106

## RESULT 12

JX0145  
hemoprotein H-450 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 29-Jan-1999  
C:Accession: JX0145  
R:Ishihara, S.; Morohashi, K.; Sadano, H.; Kawabata, S.; Gotoh, O.; Omura, T.  
J. Biochem. 108, 899-902, 1990  
A:Title: Molecular cloning and sequence analysis of cDNA coding for rat liver hemoprotein  
A:Reference number: JX0145; MUID:91210211  
A:Accession: JX0145  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-547 <ISH>  
A:Cross-references: GB:D01098; NID:g220758; PID:d1001347; PID:g220759  
C:Superfamily: cystathionine beta-synthase; CBS homology  
F:417-465/Domain: CBS homology <CBS2>

Query Match 61.8%; Score 34; DB 2; Length 547;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

## QY 2 LKCKLKLC 10

|||||  
|||:| ||

## Db 98 LKCELLAKC 106

## RESULT 13

T01377  
hypothetical protein F3D13.1 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Mar-1999  
C:Accession: T01377  
R:Corde, M.; Wollam, C.; Pape, K.  
submitted to the EMBL data Library, May 1998  
A:Description: The sequence of A. thaliana F3D13.  
A:Reference number: Z14309  
A:Accession: T01377  
A:Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA  
A:Residues: 1-991 <COR>  
A:Cross-references: EMBL:AF069300; NID:g3193305; PID:g3193306  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Introns: 450/1  
A:Note: F3D13.1

Query Match 61.8%; Score 34; DB 2; Length 991;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

## QY 1 KLKCKLKLC 9

|||||  
|||:| ||

## Db 281 KLLCRMKLC 289

## RESULT 14

I38344  
titin, cardiac muscle - human  
N:Alternate names: Connectin  
N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.1.-)  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 02-Jul-1998  
C:Accession: I38344; I38345; S20898; S20897; S20899; S63665; S37393  
R:Labait, S.; Kolmerer, B.  
Science 270, 293-296, 1995  
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.  
A:Reference number: A57430; MUID:96026330  
A:Accession: I38344  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/  
A:Molecule type: mRNA  
A:Residues: 1-26926 <LAB1>  
A:Cross-references: EMBL:X90568; NID:gl017424; PID:gl017425  
R:Musco, G.; Triatziou, C.; Schuck, P.; Pastore, A.  
Biochemistry 34, 553-561, 1995  
A:Title: Dissecting titin into its structural motifs: Identification of an alpha-heli  
A:Reference number: I38345; MUID:95119041  
A:Accession: I38345  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1977-2014 <MUS>  
A:Cross-references: EMBL:X83270; NID:g602579; PID:g602580  
A:Note: conformation and properties are reported for a synthetic peptide correspondin  
R:Labait, S.; Gautel, M.; Lakey, A.; Trinick, J.  
EMBO J. 11, 1711-1716, 1992  
A:Title: Towards a molecular understanding of titin.  
A:Reference number: S20897; MUID:92258380  
A:Accession: S20898  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 13597-14200, I', 14202-14696 <LAB2>  
A:Cross-references: EMBL:X64697; NID:g37192; PID:g37193  
A:Accession: S20897  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 16330-16382, S', 16384-16756, F', 16758-16860 <LAB3>  
A:Cross-references: EMBL:X64699; NID:g37190; PID:g37191  
A:Accession: S20899  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 22278-22431, R', 22433-22448, G', 22450-22453, Q', 22455-22480, TR', 2248

R:Kolmerer, B.; Olivieri, N.; Witt, C.C.; Herrmann, B.G.; Labait, S.  
J. Mol. Biol. 256, 556-563, 1996  
A:Title: Genomic organization of M line titin and its tissue-specific expression in t  
A:Reference number: S63665; MUID:96177761  
A:Accession: S63665  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 26729-26825 <KOL>

A:Cross-references: EMBL:X92412; NID:g1236761  
 R:Gaucel, M.; Leonard, K.; Labelit, S.  
 EMBO J. 12, 3827-3834, 1993  
 A:Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiating muscle  
 A:Reference number: S37393; MUID:94008990  
 A:Accession: S37393  
 A:Molecule type: mRNA  
 A:Residues: 26931-26926 <GAU>  
 R:Improta, S.; Politou, A.S.; Pastore, A.  
 submitted to the Brookhaven Protein Data Bank, February 1996  
 A:Reference number: A66736; PDB:1TIT  
 A:Contents: annotation; conformation by (1)H-NMR, residues 5253-5341  
 R:Pfuhl, M.; Pastore, A.  
 submitted to the Brookhaven Protein Data Bank, August 1996  
 A:Reference number: A66201; PDB:1NCT  
 A:Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155  
 C:Genetics:  
 A:Gene: GDB:TTN  
 A:Cross-references: GDB:127867; OMIM:188840  
 A:Map position: 2q31-2q32  
 C:Function:  
 A:Description: structural protein forming filaments in striated muscle  
 C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro  
 C:Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco  
 ral protein  
 F:24752-25008/Domain: protein kinase homology <KIN>  
 F:84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,  
 98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1  
 16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18  
 21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248  
 F:26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 60.0%; Score 33; DB 1; Length 26926;  
 Best Local Similarity 66.7%; Pred. NO. 3.8e+03;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLKCKLKLK 9  
 | | | | |  
 DB 2298 KTKCKLKMK 2306

RESULT 15  
 ISECAB  
 L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli  
 C:Species: Escherichia coli  
 C:Date: 30-Jun-1988 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C:Accession: F64727; A29022; S40578; S35946  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: F64727  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-500 <BLAT>  
 A:Cross-references: GB:AE000116; GB:U00096; NID:g1786240; PID:g1786248; UWGP:b0062  
 A:Experimental source: Strain K-12, substrain MG1655  
 R:Lee, N.; Gielow, W.; Martin, R.; Hamilton, E.; Fowler, A.  
 Gene 47, 231-244, 1986  
 A:Title: The organization of the arabad operon of Escherichia coli.  
 A:Reference number: A91559; MUID:87163495  
 A:Accession: A29022  
 A:Molecule type: DNA  
 A:Residues: 1-247, 'E', 249-359, 'V', 361-500 <LEE>  
 R:Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu  
 submitted to the EMBL Data Library, December 1992  
 A:Reference number: S40531  
 A:Accession: S40578  
 A:Molecule type: DNA  
 A:Residues: 1-71, 'P', 73-247, 'E', 249-359, 'V', 361-500 <YUR>

A:Cross-references: EMBL:D10483; NID:g216434; PID:d1001805; PID:g216482  
 R:Bachelier, S.; Saurin, W.; Perrin, D.; Hofnung, M.; Gilson, E.  
 submitted to the EMBL Data Library, July 1993  
 A:Description: Two major families of bacterial interspersed mosaic elements (BIWE) :  
 A:Reference number: S35946  
 A:Accession: S35946  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 458-500 <BAC>  
 A:Cross-references: EMBL:X74279; NID:g417622; PID:g417623  
 C:Genetics:  
 A:Gene: araA  
 A:Map position: 1 min  
 C:Function:  
 A:Description: catalyzes the conversion of L-arabinose to L-ribulose, the first step  
 C:Superfamily: L-arabinose isomerase  
 C:Keywords: arabinose metabolism; intramolecular oxidoreductase; isomerase

Query Match 60.0%; Score 33; DB 1; Length 500;  
 Best Local Similarity 77.8%; Pred. NO. 1.5e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLKCKLKLK 9  
 | | | | |  
 DB 43 KLPCKLVLK 51

Search completed: September 7, 1999, 23:06:30  
 Job time: 2478 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:50:18 ; Search time 71.87 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-124-280A-16

Perfect score: 55

Sequence: 1 KLKCKLKLKC 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	41	74.5	246	1 Y208_METJA	Q57661 methanococ
2	37	67.3	2068	1 T2DI_DROME	P51123 drosophila
3	35	63.6	1969	1 MYSA_CAEEL	P12844 caenorhabdi
4	34	61.8	550	1 CBS_HUMAN	P35520 homo sapien
5	34	61.8	560	1 CBS_RAT	P32232 rattus norv
6	34	61.8	417	1 Y228_CAEEL	P98061 caenorhabdi
7	33	60.0	500	1 ARAA_ECOLI	P08202 escherichia
8	33	60.0	500	1 ARAA_SALTY	P06189 salmonella
9	33	60.0	501	1 SUC1_CANAL	P33181 candida alb
10	33	60.0	215	1 TRJ9_ECOLI	Q00738 escherichia
11	32	58.2	369	1 CNA1_YEAST	P22434 saccharomyc
12	32	58.2	803	1 CUL4_CAEEL	Q17392 caenorhabdi
13	32	58.2	2411	1 DAB_DROME	P98081 drosophila
14	32	58.2	559	1 FRB7_WHEAT	Q43207 triticum ae
15	32	58.2	456	1 GSHR_HAEIN	P43783 haemophilus
16	32	58.2	478	1 GSHR_HUMAN	P00390 homo sapien
17	32	58.2	466	1 GSHR_MOUSE	P47791 mus musculu
18	32	58.2	43	1 MYX1_CROVC	P12028 crotalus vi
19	32	58.2	42	1 MYX1_CROVC	P01476 crotalus vi
20	32	58.2	43	1 MYX2_CROVC	P12029 crotalus vi
21	32	58.2	45	1 MYX2_CROVC	P19861 crotalus vi
22	32	58.2	43	1 MYXC_CROVC	P01477 crotalus vi
23	32	58.2	2616	1 NDL_DROME	P98159 drosophila
24	32	58.2	3066	1 POLG_SCHVN	Q65399 b genome po
25	32	58.2	1886	1 POL_COYMY	P19199 commelina y
26	32	58.2	639	1 PRIM_MYCTU	P95239 mycobacteri
27	32	58.2	2331	1 RRPL_MABVP	P31352 marburg vir
28	32	58.2	2331	1 RRPL_MABVP	P35262 marburg vir
29	32	58.2	273	1 Y4JE_RHISN	P55505 rhizobium s
30	32	58.2	479	1 YP66_YEAST	Q12194 saccharomyc
31	31.5	57.3	1391	1 MST2_DROHY	Q08696 drosophila
32	31	56.4	251	1 ATPD_PEA	Q02758 pisum sativ
33	31	56.4	71	1 BRJE_RANES	P32412 rana escul
34	31	56.4	1447	1 BUD4_YEAST	P47136 saccharomyc
35	31	56.4	258	1 C4BB_RAT	Q63515 rattus norv
36	31	56.4	563	1 CH60_TRYBB	Q37683 trypanosoma
37	31	56.4	562	1 CH60_TRYCR	Q95046 trypanosoma
38	31	56.4	125	1 FABL_RHASA	P98056 rhandia sap
39	31	56.4	463	1 GLPR_HUMAN	P43220 homo sapien
40	31	56.4	489	1 GLPR_MOUSE	Q35659 mus musculu
41	31	56.4	463	1 GLPR_RAT	P32301 rattus norv
42	31	56.4	1477	1 HTK7_HYDAT	Q25197 hydra atten
43	31	56.4	502	1 KS6_HUMAN	P23443 homo sapien

ALIGNMENTS

```
RESULT 1
ID Y208_METJA STANDARD; PRT; 246 AA.
AC Q57661;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0208.
GN MJ0208.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUITON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON H.-P., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKI M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
"Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC "BACTERIAL-TYPE" 4FE-4S FERREDOXINS.
CC -!- SIMILARITY: HIGH, TO M.JANNASCHII MJ0730.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67476; G1498983; -.
CC TIGR: MJ0208; -.
CC PROSITE: PS00198; 4FE4S_FERREDOXIN; 2.
CC PFAM: PF00037; fer4; 1.
CC HSP: P00208; 1BUU.
CC KW HYPOTHETICAL PROTEIN; ELECTRON TRANSPORT; IRON-SULFUR; 4FE-4S.
CC SEQUENCE 246 AA; 27696 MW; 569E2240E CRC32;
SQ
```

Query Match 74.5%; Score 41; DB 1; Length 246;  
Best Local Similarity 80.0%; Pred. No. 2.1;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KLKCKLKLKC 10  
| | | | |  
Db 156 KKKCKLKLKC 165

```
RESULT 2
T2DI_DROME STANDARD; PRT; 2068 AA.
ID T2DI_DROME
AC P51123;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TRANSCRIPTION INITIATION FACTOR TFIIID 230 KD SUBUNIT (TAFII-230)
DE (TAFII250) (TBP-ASSOCIATED FACTOR 230 KD) (P230).
```

GN TAF250.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-75 AND 540-546.  
 RX MEDLINE; 93279463.  
 RA KOKUBO T., GONG D.-W., YAMASHITA S., HORIKOSHI M., ROEDER R.G.,  
 RA NAKATANI Y.;  
 RT "Drosophila 230-kD TFIID subunit, a functional homolog of the human  
 RT cell cycle gene product, negatively regulates DNA binding of the TATA  
 RT box-binding subunit of TFIID.";  
 RL GENES DEV. 7:1033-1046(1993).  
 CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIID ASSEMBLY BY  
 CC INTERACTING WITH BOTH TBP AND OTHER TAF, AS WELL AS SERVING TO  
 CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL FOR  
 CC PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES DNA-  
 CC BINDING ACTIVITY. IS A NEGATIVE REGULATOR OF THE TATA BOX-BINDING  
 CC ACTIVITY OF TBP.  
 CC -!- SUBUNIT: TFIID IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED  
 CC FACTORS.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: CONTAINS 2 BROMODOMAINS.  
 CC -!- SIMILARITY: CONTAINS 1 HMG BOX.  
 CC -!- SIMILARITY: TO HUMAN TAFII-250 (CCG1). SOME TO S.POMBE TAFII-111  
 CC AND TO S.CEREVISIAE TAF145.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S61883; E82844; -  
 CC FLYBASE; FBgn0010355; TAF250.  
 CC PROSITE; PS00633; BROMODOMAIN\_1; 2.  
 CC PROSITE; PS0014; BROMODOMAIN\_2; 2.  
 CC PFAM; PF00439; bromodomain; 2.  
 CC TRANSFAC; T02119; -  
 CC BROMODOMAIN: NUCLEAR PROTEIN; DNA-BINDING; CELL CYCLE; REPEAT;  
 CC TRANSCRIPTION REGULATION; PHOSPHORYLATION.  
 CC FT DNA\_BIND 1247 1360 HMG BOX (POTENTIAL).  
 CC FT DOMAIN 1445 1451 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 CC FT DOMAIN 1490 1560 BROMODOMAIN.  
 CC FT DOMAIN 1612 1682 BROMODOMAIN.  
 CC FT DOMAIN 1995 2068 GLN-RICH.  
 CC FT VARIANT 575 575 P -> S.  
 CC SEQUENCE 2068 AA; 232494 MW; 225A9F98 CRC32;  
 SQ  
 Query Match 67.3%; Score 37; DB 1; Length 2068;  
 Best Local Similarity 80.0%; Pred. No. 61;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KLKCKLKLK 10  
 ||| |||||  
 Db 1342 KLKFDLKLK 1351  
 RESULT 3  
 ID MYSA\_CAEEL STANDARD; PRT; 1969 AA.  
 AC P12844;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE MYOSIN HEAVY CHAIN A (MHC A).  
 GN MYO-3.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEIA; RHABDITIA; RHABDITIDA;

OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 89178677.  
 RA DIBB N.J., MAKUYAMA I.N., KRAUSE M., KARN J.;  
 RA "Sequence analysis of the complete Caenorhabditis elegans myosin  
 RT heavy chain gene family.";  
 RL J. MOL. BIOL. 205:603-613(1989).  
 CC -!- FUNCTION: MUSCLE CONTRACTION.  
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -!- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.  
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -!- EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT MEROMYOSIN (LMM)  
 CC AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE SPLIT FURTHER INTO  
 CC 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED SUBFRAGMENT (S2).  
 CC -!- THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN C. ELEGANS.  
 CC -!- MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY WALL MUSCLE.  
 CC THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.  
 CC -!- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF  
 CC SIMILARITY WITH THE GLOBULAR HEAD SEQUENCES OF MUSCLE & NONMUSCLE  
 CC HEAVY CHAINS. BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED. BUT  
 CC THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE  
 CC THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X08067; G6799; -  
 CC PIR; S02771; S02771.  
 CC PFAM; PF00063; myosin\_head; 1.  
 CC PFAM; PF00612; IQ; 1.  
 CC HSP; P08799; LMND.  
 CC MYOSIN: MUSCLE PROTEIN; COILED COIL; THICK FILAMENT; ACTIN-BINDING;  
 CC ATP-BINDING; METHYLATION; ALKYLATION; HEPTAD REPEAT PATTERN;  
 CC MULTIGENE FAMILY.  
 CC FT DOMAIN 1 851 GLOBULAR HEAD (S1).  
 CC FT DOMAIN 852 1969 RODLIKE TAIL (S2 AND LMM DOMAINS).  
 CC FT DOMAIN 857 1969 COILED COIL (POTENTIAL).  
 CC FT NP\_BIND 179 186 ATP.  
 CC FT DOMAIN 667 689 ACTIN-BINDING.  
 CC FT DOMAIN 770 784 ACTIN-BINDING.  
 CC FT MOD\_RES 130 130 METHYLATION (TRI-) (POTENTIAL).  
 CC FT MOD\_RES 707 707 ALKYLATION (SH-1) (POTENTIAL).  
 CC FT MOD\_RES 717 717 ALKYLATION (SH-2) (POTENTIAL).  
 CC SEQUENCE 1969 AA; 225509 MW; 4E50DCA6 CRC32;  
 SQ  
 Query Match 63.6%; Score 35; DB 1; Length 1969;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLKCKLKLK 8  
 ||:|||||  
 Db 1893 KLQCKLKI 1890  
 RESULT 4  
 ID CBS\_HUMAN STANDARD; PRT; 550 AA.  
 AC P35520;  
 DT 01-JUN-1994 (REL. 29, CREATED)



DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CYSTATHIONINE BETA-SYNTHASE (EC 4.2.1.22) (SERINE SULFHYDRASE)  
 DE (BETA-THIOMASE).  
 GN CBS.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 94093551.  
 RA KRAUS J.P., LE K., SWAROOP M., OHURA T., TAHARA T., ROSENBERG L.E.,  
 RA ROPER M.D., KOZICH V.,  
 RA "Human cystathionine beta-synthase cDNA: sequence, alternative  
 RT splicing and expression in cultured cells.";  
 RL HUM. MOL. GENET. 2:1633-1638(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 95321944.  
 RA CHASSE J.-F., PALLY E., PARIS D., PAUL V., SINET P.M., KAMOUN P.,  
 RA LONDON J.;  
 RA "Genomic organization of the human cystathionine beta-synthase gene:  
 RT evidence for various cDNAs.";  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 211:826-832(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94294429.  
 RA KRUGER W.D., COX D.R.;  
 RA "A yeast system for expression of human cystathionine beta-synthase:  
 RT structural and functional conservation of the human and yeast  
 RT genes.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:6614-6618(1994).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE; 78242360.  
 RA KRAUS J.P., PACKMAN S., FOWLER B., ROSENBERG L.E.;  
 RA "Purification and properties of cystathionine beta-synthase from  
 RT human liver. Evidence for identical subunits.";  
 RL J. BIOL. CHEM. 253:6523-6528(1978).  
 RN [5]  
 RP VARIANT THR-277.  
 RX MEDLINE; 93244841.  
 RA KOZICH V., KRAUS J.P.;  
 RA "Screening for mutations by expressing patient cDNA segments in E.  
 RT coli: homocystinuria due to cystathionine beta-synthase deficiency.";  
 RL HUM. MUTAT. 1:113-123(1992).  
 RN [6]  
 RP VARIANTS VAL-113 AND LEU-144.  
 RX MEDLINE; 93357760.  
 RA KOZICH V., DE FRANCHIS R., KRAUS J.P.;  
 RA "Molecular defect in a patient with pyridoxine-responsive  
 RT homocystinuria.";  
 RL HUM. MOL. GENET. 2:815-816(1993).  
 RN [7]  
 RP VARIANTS THR-277 AND SER-306.  
 RX MEDLINE; 94108438.  
 RA HU F.L., GU Z., KOZICH V., KRAUS J.P., RAMESH V., SHIH V.E.;  
 RA "Molecular basis of cystathionine beta-synthase deficiency in  
 RT pyridoxine responsive and nonresponsive homocystinuria.";  
 RL HUM. MOL. GENET. 2:1857-1860(1993).  
 RN [8]  
 RP VARIANTS ARG-77 AND ASN-101.  
 RX MEDLINE; 95072574.  
 RA DE FRANCHIS R., KOZICH V., MCINNES R., KRAUS J.P.;  
 RT "Identical genotypes in siblings with different homocystinuric  
 RT phenotypes: identification of three mutations in cystathionine beta-  
 RT synthase using an improved bacterial expression system.";  
 RL HUM. MOL. GENET. 3:1103-1108(1994).  
 RN [9]  
 RP VARIANTS GLN-124 AND ASP-130.  
 RX MEDLINE; 9512517.

RA MARBLE M., GERAGHTY M.T., DE FRANCHIS R., KRAUS J.P., VALLE D.;  
 RT "Characterization of a cystathionine beta-synthase allele with three  
 RT mutations in cis in a patient with B6 nonresponsive homocystinuria.";  
 RL HUM. MOL. GENET. 3:1883-1886(1994).  
 RN [10]  
 RP VARIANTS  
 RX MEDLINE; 95056993.  
 RA KRAUS J.P.;  
 RT "Komrower Lecture. Molecular basis of phenotype expression in  
 RT homocystinuria.";  
 RL J. INHERIT. METAB. DIS. 17:383-390(1994).  
 RN [11]  
 RP VARIANTS SER-87; GLN-124 AND MET-256.  
 RX MEDLINE; 95282779.  
 RA SEBASTIO G., SPERANDEO M.P., PANICO M., DE FRANCHIS R., KRAUS J.P.,  
 RA ANDRIA G.;  
 RA "The molecular basis of homocystinuria due to cystathionine beta-  
 RT synthase deficiency in Italian families, and report of four novel  
 RT mutations.";  
 RL AM. J. HUM. GENET. 56:1324-1333(1995).  
 RN [12]  
 RP VARIANTS TYR-164 AND MET-370.  
 RX MEDLINE; 95362263.  
 RA KLUIJTWANS L.A.J., BLOM H.J., BOERS G.H.J., VAN OOST B.A.,  
 RA TRIJBELS F.J.M., VAN DEN HEUVEL L.P.W.J.;  
 RA "Two novel missense mutations in the cystathionine beta-synthase gene  
 RT in homocystinuric patients.";  
 RL HUM. GENET. 96:249-250(1995).  
 RN [13]  
 RP VARIANTS MET-167; HIS-223; THR-277; GLY-306; VAL-330 AND GLU-453.  
 RX MEDLINE; 96133289.  
 RA KRUGER W.D., COX D.R.;  
 RA "A yeast assay for functional detection of mutations in the human  
 RT cystathionine beta-synthase gene.";  
 RL HUM. MOL. GENET. 4:1155-1161(1995).  
 RN [14]  
 RP VARIANT LEU-289.  
 RX MEDLINE; 96057367.  
 RA SPERANDEO M.P., PANICO M., PEPE A., CANDITO M., DE FRANCHIS R.,  
 RA KRAUS J.P., ANDRIA G., SEBASTIO G.;  
 RT "Molecular analysis of patients affected by homocystinuria due to  
 RT cystathionine beta-synthase deficiency: report of a new mutation in  
 RT exon 8 and a deletion in intron 11.";  
 RL J. INHERIT. METAB. DIS. 18:211-214(1995).  
 RN [15]  
 RP VARIANTS GLU-383 AND SER-538.  
 RX MEDLINE; 97144192.  
 RA ARAL B., COUDE M., LONDON J., AUPETIT J., CHASSE J.-F., ZABOT M.-T.,  
 RA CHADEFAUX-VEREMANS B., KAMOUN P.;  
 RT "Two novel mutations (K384E and L539S) in the C-terminal moiety of  
 RT the cystathionine beta-synthase protein in two French  
 RT pyridoxine-responsive homocystinuria patients.";  
 RL HUM. MUTAT. 9:81-82(1997).  
 RN [16]  
 RP VARIANT ASN-443.  
 RX MEDLINE; 96331231.  
 RA KLUIJTWANS L.A.J., BOERS G.H.J., STEVENS E.M.B., RENIER W.O.,  
 RA KRAUS J.P., TRIJBELS F.J.M., VAN DEN HEUVEL L.P.W.J., BLOM H.J.;  
 RT "Defective cystathionine beta-synthase regulation by S-  
 RT adenosylmethionine in a partially pyridoxine responsive  
 RT homocystinuria patient.";  
 RL J. CLIN. INVEST. 98:285-289(1996).  
 RN [17]  
 RP VARIANTS LYS-143 AND TYR-164.  
 RX MEDLINE; 95072574.  
 RA GORDON R.B., COX A.J., DAWSON P.A., EMMERSON B.T., KRAUS J.P.,  
 RA DUDMAN N.P.B.;  
 RT "Mutational analysis of the cystathionine beta-synthase gene: a  
 RT splicing mutation, two missense mutations and an insertion in patients  
 RT with homocystinuria.";  
 RL HUM. MUTAT. 11:332-332(1998).  
 CC -1- CATALYTIC ACTIVITY: L-SERINE + L-HOMOCYSTEINE - CYSTATHIONINE +  
 CC H(2)O.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.

CC -!- PATHWAY: FIRST STEP IN HOMOCYSTEINE TRANSULFURATION.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- DISEASE: DEFECTS IN CBS ARE A MAJOR CAUSE OF HOMOCYSTINURIA. THE  
 CC MOST FREQUENT SYMPTOMS INCLUDE DISLOCATED OPTIC LENSES, VASCULAR  
 CC DISORDERS (ARTERIOSCLEROSIS AND THROMBOSIS), SKELETAL  
 CC ABNORMALITIES AND MENTAL RETARDATION. PATIENTS WITH HOMOCYSTINURIA  
 CC HAVE ELEVATED LEVELS OF HOMOCYSTEINE AND METHIONINE IN THEIR BODY  
 CC FLUID.  
 CC -!- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-  
 CC SYNTHASE FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 CBS DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L19501; G388716; -;  
 CC EMBL: X82166; G558582; -;  
 CC EMBL: L14577; G1289362; -;  
 CC MIM: 236200; -;  
 CC PROSITE: PS00901; CYS\_SYNTHASE; 1.  
 CC PFAM: PF00291; S\_T\_dehydratase; 1.  
 CC PFAM: PF00571; CBS; 1.  
 CC KW LYASE; CYSTEINE BIOSYNTHESIS; PYRIDOXAL PHOSPHATE; DISEASE MUTATION;  
 CC CBS DOMAIN.  
 CC FT INIT\_MET 0 0 BY SIMILARITY.  
 CC FT BINDING 118 118 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC FT DOMAIN 415 468 CBS.  
 CC FT VARIANT 77 77 P -> R (IN HOMOCYSTINURIA; 50% OF  
 CC ACTIVITY).  
 CC FT VARIANT 87 87 P -> S (IN HOMOCYSTINURIA). 50% OF  
 CC FT VARIANT 101 101 K -> N (IN HOMOCYSTINURIA; 50% OF  
 CC ACTIVITY).  
 CC FT VARIANT 113 113 A -> V (IN HOMOCYSTINURIA; 50% OF  
 CC ACTIVITY).  
 CC FT VARIANT 124 124 R -> Q (IN SEVERE HOMOCYSTINURIA;  
 CC LOSS OF ACTIVITY).  
 CC FT VARIANT 130 130 E -> D (IN HOMOCYSTINURIA; LOSS OF  
 CC ACTIVITY).  
 CC FT VARIANT 143 143 E -> K (IN HOMOCYSTINURIA).  
 CC FT VARIANT 144 144 C -> L (IN HOMOCYSTINURIA).  
 CC FT VARIANT 164 164 P -> Y (IN HOMOCYSTINURIA).  
 CC FT VARIANT 167 167 V -> M (IN HOMOCYSTINURIA).  
 CC FT VARIANT 223 223 R -> H (IN HOMOCYSTINURIA).  
 CC FT VARIANT 238 238 E -> K (IN HOMOCYSTINURIA).  
 CC FT VARIANT 256 256 T -> M (IN SEVERE HOMOCYSTINURIA;  
 CC LOSS OF ACTIVITY).

Query Match 61.8%; Score 34; DB 1; Length 550;  
 Best Local Similarity 66.7%; Pred. No. 60;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLKC 10  
 |||:| ||  
 Db 100 LKCELLAKC 108

RESULT 5  
 CBS\_RAT STANDARD; PRT; 560 AA.  
 AC P32232;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CYSTATHIONINE BETA-SYNTHASE (EC 4.2.1.22) (SERINE SULFHYDRASE)  
 DE (BETA-THIONASE) (HEMOPROTEIN H-450).  
 GN CBS.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 CC [1]  
 CC RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 CC RP TISSUE=LIVER;  
 CC RX MEDLINE; 92283859.  
 CC RA SWAROOP M., BRADLEY K., OHURA T., TAHARA T., ROPER M.D.,  
 CC ROSENBERG L.E., KRAUS J.P.;  
 CC RT "Rat cystathionine beta-synthase. Gene organization and alternative  
 CC splicing.";  
 CC RL J. BIOL. CHEM. 267:11455-11461(1992).  
 CC [2]  
 CC RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-22 AND 38-47.  
 CC RP STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 CC RX MEDLINE; 91210211.  
 CC RA ISHITHARA S., MOROHASHI K.-I., SADANO H., KAWABATA S.-I., GOTOH O.,  
 CC OMURA T.;  
 CC RT "Molecular cloning and sequence analysis of cDNA coding for rat liver  
 CC hemoprotein H-450.";  
 CC RL J. BIOCHEM. 108:899-902(1990).  
 CC -!- CATALYTIC ACTIVITY: L-SERINE + L-HOMOCYSTEINE = CYSTATHIONINE +  
 CC H(2O).  
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -!- PATHWAY: FIRST STEP IN HOMOCYSTEINE TRANSULFURATION.  
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, AND BRAIN.  
 CC -!- PTM: BINDS COVALENTLY TO A HEME GROUP THROUGH A THIOLATE  
 CC LIGAND.  
 CC -!- ALTERNATIVE PRODUCTS: FOUR VARIANTS CAN BE FORMED BY ALTERNATIVE  
 CC SPLICING (TYPE I THROUGH IV). THE SEQUENCE OF TYPE I IS SHOWN.  
 CC -!- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-  
 CC SYNTHASE FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 CBS DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M88344; G206597; -;  
 CC EMBL: M88346; G206600; -;  
 CC EMBL: D01098; D1001347; -;  
 CC PIR: JX0145; JX0145.  
 CC PIR: A42790; A42790.  
 CC PROSITE: PS00901; CYS\_SYNTHASE; 1.  
 CC PFAM: PF00291; S\_T\_dehydratase; 1.  
 CC PFAM: PF00571; CBS; 1.  
 CC KW LYASE; CYSTEINE BIOSYNTHESIS; PYRIDOXAL PHOSPHATE; HEME; CBS DOMAIN;  
 CC ALTERNATIVE SPLICING.  
 CC FT INIT\_MET 0 0  
 CC FT BINDING 115 115 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC FT DOMAIN 411 464 CBS.  
 CC FT VARSPLIC 513 513 S -> Y (IN TYPE III).  
 CC FT VARSPLIC 514 527 MISSING (IN TYPE III).  
 CC FT VARSPLIC 449 477 ALLGNVTGLGNMISLLACKVRPSDEVCKV ->  
 CC LRQSKDICHPTKRHIQAHGLRKKVPTDEA (IN TYPE  
 CC IV).  
 CC FT VARSPLIC 478 560 MISSING (IN TYPE III).  
 CC FT CONFLICT 414 414 L -> P (IN REF. 2).  
 CC SQ SEQUENCE 560 AA; 61323 MW; 6D972105 CRC32;

Query Match 61.8%; Score 34; DB 1; Length 560;  
 Best Local Similarity 66.7%; Pred. No. 61;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKCKLKLKC 10  
 |||:| ||  
 Db 97 LKCELLAKC 105

```
RESULT 6
YP28_CAEEL STANDARD; PRT; 417 AA.
AC P98061;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL ZINC METALLOPROTEINASE F42A10.8 PRECURSOR (EC 3.4.24.-).
GN F42A10.8.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA LATRILLE P.;
RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC
METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10414; G500741; -
CC WORPEP; F42A10.8; CE01299.
CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
CC DR PROSITE; PS01180; CUB; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE_NEG.
CC DR PFAM; PF00099; zn-protease; 1.
CC DR PFAM; PF00431; CUB; 1.
CC DR HSP; P28825; LIAP.
CC KW HYPOTHETICAL PROTEIN; REPEAT: HYDROLASE; METALLOPROTEASE;
KW EGF-LIKE DOMAIN; ZINC; SIGNAL.
FT SIGNAL 1 14
FT CHAIN 15 417
FT METAL 134 134
FT ACT_SITE 135 135
FT METAL 138 138
FT METAL 144 144
FT REPEAT ? ?
FT DOMAIN 284 403
FT CARBOHYD 76 76
FT CARBOHYD 237 237
FT CARBOHYD 314 314
FT SEQUENCE 417 AA; 46790 MW; 4217002C CRC32;
SQ
Query Match 61.8%; Score 34; DB 1; Length 417;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 3 KCKLKLKC 10
Db 241 QCKIQMKC 248
:|:|:|:|
RESULT 7
ARAA_ECOLI STANDARD; PRT; 500 AA.
AC P08202; P78040;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
```

```
DE L-ARABINOSE ISOMERASE (EC 5.3.1.4).
GN ARAA.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87163495.
RA LEE N., GIELOW W., MARTIN R., HAMILTON E., FOWLER A.;
RT "The organization of the arabad operon of Escherichia coli.";
RL GENE 47:231-244(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 92334977.
RA YURA T., MORI H., NAGAI H., NAGATA T., ISHIHAMA A., FUJITA N.,
RA ISONO K., MIZOBUCHI K., NAKATA A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
the 0-2.4 min region.";
RL NUCLEIC ACIDS RES. 20:3305-3308(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474(1997).
RN [4]
RP SEQUENCE OF 458-500 FROM N.A.
RA BACHELIER S., SAURIN W., PERRIN D., HOFNUNG M., GILSON E.;
RL SUBMITTED (JUL-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: L-ARABINOSE -> L-RIBULOSE.
CC -!- PATHWAY: FIRST STEP OF L-ARABINOSE CATABOLISM AFTER IT IS
TRANSPORTED INTO THE CELL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15263; G145305; -
CC EMBL; D10483; G216482; -
CC EMBL; AF000116; G1786248; -
CC EMBL; X74279; G417623; -
CC PIR; A29022; ISECAB.
CC PIR; S35946; S35946.
CC PIR; S40578; S40578.
CC ECODBASE; G054.1; 6TH EDITION.
CC ECOGENE; EG10052; ARAA.
KW ARABINOSE CATABOLISM; ISOMERASE.
FT CONFLICT 72 72
FT CONFLICT 248 248
FT CONFLICT 360 360
FT SEQUENCE 500 AA; 56103 MW; 60A73BCF CRC32;
SQ
Query Match 60.0%; Score 33; DB 1; Length 500;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 KLKCKLKLK 9
Db 43 KLPCKLVLK 51
||| ||| |||
RESULT 8
ARAA_SALTY
```

```

ID  ARAA_SALTY      STANDARD;      PRT;      500 AA.
AC  P06189;
DT  01-JAN-1988 (REL. 06, CREATED)
DT  01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT  01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE  L-ARABINOSE ISOMERASE (EC 5.3.1.4).
GN  ARAA.
OS  SALMONELLA TYPHIMURIUM.
OC  BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC  SALMONELLA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=LT2;
RX  MEDLINE; 85232045.
RA  LIN H.-C., LEI S.-P., WILCOX G.;
RT  "The arabid operon of Salmonella typhimurium LT2. II. Nucleotide
RT  sequence of araa and primary structure of its product, L-arabinose
RT  isomerase."
RL  GENE 34:123-128(1985).
CC  -!- CATALYTIC ACTIVITY: L-ARABINOSE = L-RIBULOSE.
CC  -!- PATHWAY: FIRST STEP OF L-ARABINOSE CATABOLISM AFTER IT IS
CC  TRANSPORTED INTO THE CELL.
CC  -!- PTM: THE N-TERMINAL IS PROBABLY MODIFIED.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M1047; G153868; -
DR  PIR; A24985; ISEBAB.
DR  STYGENE; SG10012; ARAA.
KW  ISOMERASE; ARABINOSE CATABOLISM.
SQ  SEQUENCE 500 AA; 55877 MW; 7F6018E0 CRC32;

Query Match      60.0%; Score 33; DB 1; Length 500;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 KLCKLKLK 9
Db  43 KLCKLVLK 51

RESULT 9
ID  SUC1_CANAL      STANDARD;      PRT;      501 AA.
AC  P33181;
DT  01-OCT-1993 (REL. 27, CREATED)
DT  01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT  01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE  PROBABLE SUCROSE UTILIZATION PROTEIN SUC1.
GN  SUC1.
OS  CANDIDA ALBICANS (YEAST).
OC  EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC  CANDIDACEAE; CANDIDA.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 92104966.
RA  KELLY R., KWON-CHUNG K.J.;
RT  "A zinc finger protein from Candida albicans is involved in sucrose
RT  utilization."
RL  J. BACTERIOL. 174:222-232(1992).
CC  -!- FUNCTION: AFFECTS SUCROSE UTILIZATION AND ALPHA-GLUCOSIDASE
CC  ACTIVITY. PROBABLE TRANSCRIPTIONAL ACTIVATOR.
CC  -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC  -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC  CLUSTER DOMAIN.
CC  -----

```

```

CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; S75352; E91548; -
DR  PIR; A43302; A43302.
DR  PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR  PROSITE; PS00448; ZN2_CY6_FUNGAL_2; 1.
DR  PFAM; PF00172; Zn_c1us; 1.
KW  TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;
KW  ZINC; METAL-BINDING.
FT  DNA_BIND 13 39      ZN(2)-CYS(6), FUNGAL-TYPE.
SQ  SEQUENCE 501 AA; 57110 MW; BA45EA95 CRC32;

Query Match      60.0%; Score 33; DB 1; Length 501;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  1 KLCKLKLK 10
Db  20 KVKCDMKTPC 29

RESULT 10
ID  TRJ9_ECOLI      STANDARD;      PRT;      215 AA.
AC  Q00738;
DT  01-APR-1993 (REL. 25, CREATED)
DT  01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT  01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE  TRAJ PROTEIN.
GN  TRAJ.
OS  ESCHERICHIA COLI.
OC  PLASMID INCFI P307.
OC  BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC  ESCHERICHIA.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=711;
RX  MEDLINE; 91261994.
RA  GRAUS-GOELDNER A., GRAUS H., SCHLACHER T., HOEGENAUER G.;
RT  "The sequences of genes bordering orit in the enterotoxin plasmid
RT  P307: comparison with the sequences of plasmids F and RL."
RL  PLASMID 24:119-131(1990).
CC  -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR POSITIVELY REGULATING THE
CC  EXPRESSION OF TRANSFER GENES THAT ARE INVOLVED IN THE CONJUGAL
CC  TRANSFER OF DNA BETWEEN BACTERIAL CELLS.
CC  -!- SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M62986; G150466; -
DR  PIR; C37390; C37390.
KW  OUTER MEMBRANE; CONJUGATION; TRANSCRIPTION REGULATION; ACTIVATOR;
KW  PLASMID.
SQ  SEQUENCE 215 AA; 25028 MW; F71CAEFE CRC32;

Query Match      60.0%; Score 33; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 KLKCKLKL 8
DB 67 KLECKLOL 74

RESULT 11
CNAL_YEAST STANDARD; PRT; 369 AA.
AC P22434;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE 3', 5'-CYCLIC-NUCLEOTIDE PHOSPHODIESTERASE 1 (EC 3.1.4.17) (PDEASE 1)
DE (LOW-AFFINITY CAMP PHOSPHODIESTERASE).
GN PDEL OR YGL248W OR NRB369.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88065501.
RA NIKAWA J.-I., SASS P., WIGLER M.;
RT "Cloning and characterization of the low-affinity cyclic AMP
phosphodiesterase gene of Saccharomyces cerevisiae.";
RL MOL. CELL. BIOL. 7:3629-3636(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE: 97127827.
RA COISSAC E., MAILLIER E., ROBINEAU S., NETTER P.;
RT "Sequence of a 39,411 bp DNA fragment covering the left end of
chromosome VII of Saccharomyces cerevisiae.";
RL YEAST 12:1555-1562(1996).
CC -1- FUNCTION: CONTROLS THE LEVEL OF CAMP IN YEAST CELLS, TOGETHER WITH
THE HIGH-AFFINITY CAMP PHOSPHODIESTERASE (PDE2).
CC -1- CATALYTIC ACTIVITY: NUCLEOSIDE 3',5'-CYCLIC PHOSPHATE + H(2)O -
NUCLEOSIDE 5'-PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO CAMP PHOSPHODIESTERASES CLASS-II.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17781; G172229; -
DR EMBL: X94337; E215623; -
DR EMBL: 272770; E243903; -
DR PIR: S05879; ESBYPC.
DR SGD: L0001358; PDEL.
DR PROSITE: PS00607; PDEASE_II; 1.
KW HYDROLASE; CAMP.
FT CONFLICT 94 L -> F (IN REF. 1).
SQ SEQUENCE 369 AA; 42016 MW; B756214A CRC32;

Query Match 58.2%; Score 32; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KLKCLK 10
DB 185 KLKCLK 190

RESULT 12
CUL4_CAEEL STANDARD; PRT; 803 AA.
AC Q17392; Q20428;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

Query Match 58.2%; Score 32; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KLKCLK 10
DB 185 KLKCLK 190

RESULT 12
CUL4_CAEEL STANDARD; PRT; 803 AA.
AC Q17392; Q20428;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
```

```
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CUL-4 PROTEIN.
GN CUL-4 OR F45E12.3.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 96279828.
RA KIPREOS E.T., LANDER L.E., WING J.P., HE W.W., HEDGECOCK E.M.;
RT "cul-1 is required for cell cycle exit in C. elegans and identifies a
novel gene family.";
RL CELL 85:829-839(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA LATREILLE P.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- DEVELOPMENTAL STAGE: HIGHEST LEVELS IN EMBRYOS AND LOWER LEVELS IN
LARVAE AND ADULTS.
CC -1- SIMILARITY: BELONGS TO THE CULLIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U58086; G1381140; -
DR EMBL: U29536; G868263; -
DR WORMPEP: F45E12.3; CE02739.
DR PROSITE: PS01256; CULLIN_1; 1.
DR PROSITE: PSS00069; CULLIN_2; 1.
DR PFAM: PF00888; Cullin; 1.
FT CONFLICT 506 515 EKMVLCKLKT -> GTD (IN REF. 2).
SQ SEQUENCE 803 AA; 92592 MW; 41FAD312 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 803;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CKLKLKC 10
DB 511 CKLKLKC 517

RESULT 13
DAB_DROME STANDARD; PRT; 2411 AA.
AC P98081;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE DISABLED PROTEIN.
GN DAB.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93194063.
RA GERTLER F.B., HILL K.K., CLARK M.J., HOFFMANN F.M.;
RT "Dosage-sensitive modifiers of Drosophila abl tyrosine kinase
function: prospero, a regulator of axonal outgrowth, and disabled, a
novel tyrosine kinase substrate.";
RL GENES DEV. 7:441-453(1993).
CC -1- FUNCTION: TOGETHER WITH ABL INVOLVED IN EMBRYONIC NEURAL
DEVELOPMENT.
CC -----
```

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- TISSUE SPECIFICITY: CNS AXONS AND BODY WALL MUSCLES.  
 CC -!- DEVELOPMENTAL STAGE: EMBRYONIC AXONEMOGENESIS.  
 CC -!- PTM: PROBABLY PHOSPHORYLATED BY THE ABL TYROSINE KINASE.  
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED FROM THE SAME GENE  
 CC BY ALTERNATIVE SPLICING.  
 CC -!- SIMILARITY: CONTAINS 1 PID DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L08845; G1498252; ALT\_SEQ.  
 DR PIR: A46299; A46299.  
 DR FLYBASE: FBgn0000414; Dab.  
 DR PROSITE: PS01179; PID; 1.  
 DR PFAM: PF00640; PID; 1.  
 DR ALTERNATIVE SPLICING: PHOSPHORYLATION.  
 FT DOMAIN 46 196  
 FT VARSPIC 462 673 MISSING (IN SHORTER FORM).  
 FT DOMAIN 1689 1801 REPEAT-RICH REGION.  
 FT REPEAT 1689 1700 ALTERNATE ARG AND ACIDIC RESIDUE.  
 FT REPEAT 1740 1750 ALTERNATE ARG AND ACIDIC RESIDUE.  
 FT REPEAT 1791 1801 ALTERNATE ARG AND ACIDIC RESIDUE.  
 FT MOD\_RES 111 111 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 482 482 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1662 1662 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1667 1667 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1701 1701 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1704 1704 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1713 1713 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1739 1739 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1826 1826 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 FT MOD\_RES 1961 1961 PHOSPHORYLATION (BY ABL) (PROBABLE).  
 SQ SEQUENCE 2411 AA; 264046 MW; 4EC2EDA7 CRC32;

Query Match 58.2%; Score 32; DB 1; Length 2411;  
 Best Local Similarity 66.7%; Pred. No. 4.6e+02;  
 Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 KLCKKLK-LKC 10  
 Db 2384 KLCKLDDFLKC 2395  
 ||||| |||

RESULT 14  
 ID FKX7\_WHEAT STANDARD; PRT; 559 AA.  
 AC Q43207;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE 70 KD PEPTIDYLPROLYL ISOMERASE (EC 5.2.1.8) (PEPTIDYLPROLYL CIS-TRANS  
 DE ISOMERASE) (CYCLOPHILIN) (PPIASE).  
 GN FKBP70.  
 OS TRITICUM AESTIVUM (WHEAT).  
 OS EUKARYOTA: VIRIDIPHYTES: EMBRYOPHYTES: TRACHEOPHYTES;  
 OC EUKARYOTA: VIRIDIPHYTES: SPERMATOPHYTES; MAGNOLIOPHYTES; LILIOPSIDA; POALES;  
 OC POACEAE; TRITICUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. ATIR; TISSUE=ROOT TIP;  
 RA OSHRA B., BREIMAN A.;  
 RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: PPIASE THAT BINDS CALMODULIN.  
 CC -!- CATALYTIC ACTIVITY: PEPTIDYLPROLINE (OMEGA=180) -  
 CC PEPTIDYLPROLINE (OMEGA=0).  
 CC -!- INDUCTION: BY HEAT SHOCK.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, TO FKBP-TYPE PPIASES.  
 CC CONTAINS THREE FKBP-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X86903; G854626; .  
 DR PROSITE: PS00453; FKBP\_PPIASE\_1; 1.  
 DR PROSITE: PS00454; FKBP\_PPIASE\_2; 3.  
 DR PROSITE: PS00059; FKBP\_PPIASE\_3; 3.  
 DR PFAM: PF00254; FKBP; 2.  
 DR HSP: P27124; IROU.  
 DR ISOMERASE; ROTAMASE; REPEAT; HEAT SHOCK; CALMODULIN-BINDING.  
 KW DOMAIN 60 148 PPIASE, FKBP-TYPE 1.  
 FT DOMAIN 176 265 PPIASE, FKBP-TYPE 2.  
 FT DOMAIN 293 384 PPIASE, FKBP-TYPE 3.  
 SQ SEQUENCE 559 AA; 62056 MW; 58793D58 CRC32;  
 Query Match 58.2%; Score 32; DB 1; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 CKLKLK 9  
 Db 459 CKLKLK 464  
 |||||  
 RESULT 15  
 GSHR\_HAEIN STANDARD; PRT; 456 AA.  
 ID GSHR\_HAEIN AC P43783;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GLUTATHIONE REDUCTASE (EC 1.6.4.2) (GR) (GRASE).  
 GN GOR OR H10161.  
 OS HAEMOPHILUS INFLUENZAE.  
 OS BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;  
 OC HAEMOPHILUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20;  
 RA BARCAK G.J., HEIMER S.R.;  
 RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20;  
 RX MEDLINE; 95350630.  
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
 RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
 RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
 RA FINE L.D., FRITZMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,  
 RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL SCIENCE 269:496-512(1995).  
 CC -!- FUNCTION: MAINTAIN HIGH LEVELS OF REDUCED GLUTATHIONE IN THE  
 CC CYTOSOL (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: NADPH + OXIDIZED GLUTATHIONE = NADP(+) +  
 CC 2 GLUTATHIONE.  
 CC -!- COFACTOR: FAD (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

```
CC -!- THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20229; G644854; -
DR EMBL; U32702; G1573119; -.
DR TIGR; H10161; -.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
DR PFAM; PF00070; pyr_redox; 1.
DR HSP; P00390; 1ALG
KW REDOX-ACTIVE CENTER; OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NADP.
FT NP_BIND 6 36 FAD (ADP PART) (PROBABLE).
FT DISULFID 42 47 REDOX-ACTIVE (BY SIMILARITY).
FT NP_BIND 298 308 FAD (FLAVIN PART) (BY SIMILARITY).
FT ACT_SITE 445 445 BY SIMILARITY.
SQ SEQUENCE 456 AA; 49329 MW; 9A507FF6 CRC32;
```

```
Query Match 58.2%; Score 32; DB 1; Length 456;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CKLKLKC 10
DB 395 CRMKLVC 401
```

Search completed: September 7, 1999, 23:50:19  
Job time: 1958 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:16 ; Search time 71.09 Seconds  
(without alignments)  
4.374 Million cell updates/sec

Title: US-09-124-280A-37  
Perfect score: 54  
Sequence: 1 IKFLKFLKFLK 11  
Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	70.4	394	1 ALR2_BACSU	P94494 bacillus su
2	36	66.7	956	1 PMAL_LYCES	P22180 lycopersico
3	35	64.8	979	1 P115_MYCHR	P41508 mycoplasma
4	34.5	63.9	238	1 Y037_METJA	O60344 methanococ
5	34	63.0	583	1 CRT1_PHYBL	P54982 phycomyces
6	34	63.0	578	1 RSP2_YEAST	P50109 saccharomyc
7	34	63.0	103	1 RR6_CYACA	O19917 cyanidium c
8	34	63.0	337	1 RT07_ACACA	P46756 acanthamoeb
9	34	63.0	153	1 Y587_METJA	O58007 methanococ
10	33	61.1	140	1 FLAE_METJA	O58306 methanococ
11	33	61.1	2291	1 RRPB_BEV	P18458 berne virus
12	33	61.1	1858	1 VGNB_CPSMV	P36312 cowpea seve
13	33	61.1	1659	1 VIT_ONCNY	Q92093 oncorhynchu
14	33	61.1	1122	1 YG3C_YEAST	P53280 saccharomyc
15	33	61.1	558	1 YG2B_YEAST	P53059 saccharomyc
16	33	61.1	110	1 YJ20_YEAST	P47092 saccharomyc
17	33	61.1	138	1 YYAN_BACSU	P37510 bacillus su
18	32.5	60.2	585	1 BLAR_STAAR	P18357 staphylococ
19	32.5	60.2	811	1 RFXL_YEAST	P48743 saccharomyc
20	32	59.3	943	1 DIP2_YEAST	Q12220 saccharomyc
21	32	59.3	190	1 NU5M_ARBLL	Q33753 arabacia lix
22	32	59.3	201	1 NU6M_CYACA	P48925 cyanidium c
23	32	59.3	235	1 PEXB_YEAST	Q12462 saccharomyc
24	32	59.3	1341	1 RPA1_METJA	Q58445 methanococ
25	32	59.3	1580	1 SUR_HUMAN	Q09428 homo sapien
26	32	59.3	422	1 SVY1_BACSU	P22326 bacillus su
27	32	59.3	1866	1 VGNB_CPMV	P03600 cowpea mosa
28	32	59.3	119	1 VSDE_SALDU	P24421 salmonella
29	32	59.3	346	1 Y343_MYCGE	P7585 mycoplasma
30	32	59.3	342	1 YM12_YEAST	Q04215 saccharomyc
31	32	59.3	824	1 Y015_CAEEL	P34651 caenorhabdi
32	31	57.4	812	1 EBN3_EBV	P12977 Epstein-Bar
33	31	57.4	762	1 MSU2_HELPY	Q25338 helicobacte
34	31	57.4	869	1 PLCL1_YEAST	P32383 saccharomyc
35	31	57.4	2182	1 POLG_COXBI	P08291 coxsackievi
36	31	57.4	2201	1 POLG_CXA9	P21404 coxsackievi
37	31	57.4	1374	1 POLG_EC11G	P29813 echovirus 1
38	31	57.4	2185	1 POLG_SCDVH	P16604 s genome po
39	31	57.4	2185	1 POLG_SVDVU	P13900 s genome po
40	31	57.4	230	1 R1BR_BACSU	P94465 bacillus su
41	31	57.4	179	1 RL15_ARCFU	O28376 haemaglob
42	31	57.4	2210	1 RRPO_TACY	P20430 tacaribe vi
43	31	57.4	293	1 RS3_BORBU	P94273 borrelia bu

ALIGNMENTS

```
RESULT 1
ALR2_BACSU
ID ALR2_BACSU STANDARD; PRT; 394 AA.
AC P94494;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PUTATIVE ALANINE RACEMASE (EC 5.1.1.1).
GN YNCD.
OS BACILLUS SUBTILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RA BORCHERT S., KLEIN C., PIKSA B., HAMMELMANN M., ENTIAN K.D.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: L-ALANINE -> D-ALANINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U66480; G1750128; -.
CC EMBL; 299113; E1183423; -.
CC DR SUBTILIST; BG12267; YNCD.
CC DR PROSITE; PS00395; ALANINE_RACEMASE; 1.
CC DR PFAM; PF00842; Ala_racemase; 1.
CC DR HYPOTHETICAL PROTEIN; ISOMERASE; PYRIDOXAL PHOSPHATE.
CC FT BINDING 39 39 PYRIDOXAL PHOSPHATE.
CC SQ SEQUENCE 394 AA; 43648 MW; B5C38A51 CRC32;

Query Match 70.4%; Score 38; DB 1; Length 394;
Best Local Similarity 70.0%; Pred. No. 5.8;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11
   ||: || |||
DB 187 KFISFLSFLK 196

RESULT 2
PMAL_LYCES
ID PMAL_LYCES STANDARD; PRT; 956 AA.
AC P22180;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).
GN LHA1
OS LYCOPERSICON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPALTA; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; SOLANACEAE; SOLANACEAE; SOLANACEAE; SOLANACEAE;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. CASTELMART; TISSUE-ROOT;
RC EWING N.N., WIMMERS L.E., MEYER D.J., CHETELAT R.T., BENNETT A.B.;
RT "Molecular cloning of tomato plasma membrane H+-ATPase.";
RL PLANT PHYSIOL. 94:1874-1881(1990).
```

CC -!- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A  
 CC HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE  
 CC ACTIVE TRANSPORT OF NUTRIENTS BY H<sup>+</sup>-SYMPORT. THE RESULTING  
 CC EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKALIZATION MAY MEDIATE  
 CC GROWTH RESPONSES.  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.  
 CC -!- SUBUNIT: POSSIBLY EXISTS AS AN HOMODIMER OR AN HOMOTRIMER.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- AS MANY AS 6 TO 8 CLOSELY RELATED GENES MAY ENCODE OTHER ISOFORMS  
 CC OF PLASMA MEMBRANE ATPASE IN TOMATO, LIKE THE LHA2 GENE PRODUCT  
 CC WHICH IS 96% IDENTICAL TO THE LHA1 GENE PRODUCT.  
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
 CC (EI-E2 ATPASES).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M60166; G170464; -  
 CC PIR: A45506; A45506  
 CC PROSITE: PS00154; ATPASE\_EI\_E2: 1.  
 CC PFAM: PF00122; EI-E2\_ATPase: 1.  
 CC KW HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;  
 CC ATP-BINDING; MULTIGENE FAMILY.  
 CC -----  
 CC DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 65 85 1 (POTENTIAL).  
 CC DOMAIN 86 96 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 97 116 2 (POTENTIAL).  
 CC DOMAIN 117 245 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 246 267 3 (POTENTIAL).  
 CC DOMAIN 268 279 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 280 302 4 (POTENTIAL).  
 CC DOMAIN 303 677 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 678 696 5 (POTENTIAL).  
 CC DOMAIN 697 712 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 713 736 6 (POTENTIAL).  
 CC DOMAIN 737 794 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 795 815 7 (POTENTIAL).  
 CC DOMAIN 816 830 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 831 847 8 (POTENTIAL).  
 CC DOMAIN 848 956 EXTRACELLULAR (POTENTIAL).  
 CC MOD\_RES 333 333 PHOSPHORYLATION (BY SIMILARITY).  
 CC BINDING 427 427 ATP (BY SIMILARITY).  
 CC SEQUENCE 956 AA; 105103 MW; E798A2DA CRC32;

Query Match 66.7%; Score 36; DB 1; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 30;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 DB 61 KFLKFLGFM 69  
 ||||| |

RESULT 3  
 ID P115\_MYCHR STANDARD; PRT; 979 AA.  
 AC P41508;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE P115 PROTEIN.  
 OS MYCOPLASMA HYORHINIS.  
 CC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;  
 CC MYCOPLASMATACEAE; MYCOPLASMA.  
 CC [1]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE; 91138990.

RA NOTARNICOLA S.M., MCINTOSH M.A., WISE K.S.;  
 RT "A mycoplasma hyorhinis protein with sequence similarities to  
 RL nucleotide-binding enzymes."; GENE 97:77-85(1991).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS  
 CC FLANKED BY PUTATIVE GLOBAL REGIONS AT THE N- AND C-TERMINUS.  
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA  
 CC P115.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M34956; G150165; -  
 CC PIR: JQ0894; JQ0894.  
 CC KW ATP-BINDING; COILED COIL.  
 CC FT NP\_BIND 32 39 ATP (POTENTIAL).  
 CC FT DOMAIN 169 224 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 231 400 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 569 821 COILED COIL (POTENTIAL).  
 CC FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).  
 CC SEQUENCE 979 AA; 110566 MW; B51BE656 CRC32;

Query Match 64.8%; Score 35; DB 1; Length 979;  
 Best Local Similarity 54.5%; Pred. No. 46;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
 DB 918 IRYEFLKLLK 928  
 |:::|||||

RESULT 4  
 ID Y037\_METJA STANDARD; PRT; 238 AA.  
 AC Q60344;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MJ0037.  
 GN MJ0037.  
 OS METHANOCOCCUS JANNASCHII.  
 CC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEORGEHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON H.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 Jannaschii".  
 RL JANNASCHII.  
 CC SCIENCE 273:1058-1073(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```
CC -----
DR EMBL; U67462; G1498797; -.
DR TIGR; M30037; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 238 AA; 26884 MW; C5CBD35C CRC32;

Query Match      63.9%; Score 34.5; DB 1; Length 238;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 1 IKFLK-FLAFLK 11
   ||||| |::||:
DB 85 IKFLKEFIEFLR 96

RESULT 5
CRTI.PHYBL
ID CRTI.PHYBL STANDARD; PRT; 583 AA.
AC P54362;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PHYTOENE DEHYDROGENASE (EC 1.3.-.-) (PHYTOENE DESATURASE).
GN CARB.
OS PHYCOMYCES BLAKESLEEANUS.
OC EUKARYOTA; FUNGI; ZYGOMYCOTA; ZYGOMYCETES; MUCORALES; MUCORACEAE;
OC PHYCOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 1555;
RA RUI2-HIDALGO M.J.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
CC -!- COFACTOR: FAD FLAVOPROTEIN (PROBABLE).
CC -!- PATHWAY: CAROTENOID BIOSYNTHESIS.
CC -!- SIMILARITY: TO BACTERIAL PHYTOENE DEHYDROGENASES AND TO BACTERIAL
CC METHOXYNEUROSPORENE DEHYDROGENASE (CRTD).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X78434; G468557; -.
DR PROSITE; P500982; PHYTOENE_DH; 1.
KW CAROTENOID BIOSYNTHESIS; OXIDOREDUCTASE; FAD; FLAVOPROTEIN; NAD;
KW TRANSMEMBRANE.
FT NP_BIND 8 41 FAD (ADP PART) (POTENTIAL).
FT TRANSMEM 531 551 POTENTIAL.
SQ SEQUENCE 583 AA; 65983 MW; 78EDB4F CRC32;

Query Match      63.0%; Score 34; DB 1; Length 583;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 FLKFLAFLK 11
   ||::| |::|
DB 122 FLRLDFMK 130

RESULT 6
PSP2_YEAST
ID PSP2_YEAST STANDARD; PRT; 578 AA.
AC P50109; Q07170;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
CC -----
```

```
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PSP2 PROTEIN (MITOCHONDRIAL REGULATOR OF SPLICING 15).
DE PSP2 OR MRS15 OR YML017W OR YM9571.01.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA FORMOSA T.D., NITTIS T.;
RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-DBI747;
RA TEPLY R.;
RL THESIS (1995), VIENNA BIOCENTRE, AUSTRIA.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA GENTILES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: DNA POLYMERASE ALPHA MUTATION SUPPRESSOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33116; G995753; -.
DR EMBL; U29398; G1808701; -.
DR EMBL; Z49810; G854473; -.
DR SGD; L0002882; PSP2.
FT DOMAIN 393 402 POLY-ASN.
FT DOMAIN 438 453 POLY-ASN.
FT DOMAIN 457 460 POLY-ASN.
SQ SEQUENCE 578 AA; 64020 MW; 7DA753A1 CRC32;

Query Match      63.0%; Score 34; DB 1; Length 578;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKF 9
   ||::| |::|
DB 117 KTKFKFKF 124

RESULT 7
RR6_CYACA
ID RR6_CYACA STANDARD; PRT; 103 AA.
AC O19917;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S6.
GN RPS6.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OG CHLOROPLAST.
OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDACEAE;
OC CYANIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK-1;
RA GLOECKNER G., ROSENTHAL A., VALENTIN K.;
RL "Organisation of 46 kb of the Cyanidium caldarium Rk1 plastid
RL genome."
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: BINDS TOGETHER WITH S18 TO 16S RIBOSOMAL RNA
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF02186; G2465743; -  
 CC DR PROSITE: PS01048; RIBOSOMAL\_S6; FALSE\_NEG.  
 CC DR PFAM: PF01250; RIBOSOMAL\_S6; 1.  
 CC KW RIBOSOMAL PROTEIN; RNA-BINDING; CHLOROPLAST.  
 CC SQ SEQUENCE 103 AA; 12340 MW; CC91D65A CRC32;

Query Match 63.0%; Score 34; DB 1; Length 103;  
 Best Local Similarity 70.0%; Pred. No. 7.8;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
 ||||| :|  
 Db 28 KFLKFLSEIK 37

RESULT 8  
 RT07\_ACACA STANDARD; PRT; 337 AA.  
 AC P46756;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.  
 GN RPS7.  
 OS ACANTHAMOEBA CASTELLANI (AMOEB).  
 OC MITOCHONDRION.  
 OG EUKARYOTA; ACANTHAMOEBAEIDAE; ACANTHAMOEBA.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30010 / NEFF;  
 RX MEDLINE: 93157849.  
 RA LONERGAN K.M., GRAY M.W.;  
 RA "Editing of transfer RNAs in Acanthamoeba castellanii mitochondria."  
 RL SCIENCE 259:812-816(1993).  
 CC -|- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
 CC -|- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U12386; G562050; -  
 CC DR PROSITE: PS00052; RIBOSOMAL\_S7; FALSE\_NEG.  
 CC DR PFAM: PF00177; S7; 1.  
 CC KW RIBOSOMAL PROTEIN; MITOCHONDRION.  
 CC SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 63.0%; Score 34; DB 1; Length 337;  
 Best Local Similarity 55.6%; Pred. No. 25;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 FLKFLKFLK 11  
 :||:|  
 Db 326 YIFLFLK 334

RESULT 9  
 Y587\_METJA STANDARD; PRT; 153 AA.  
 ID Y587\_METJA

AC Q58007;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MJ0587.  
 GN MJ0587.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE: 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., FUHRMANN J.L., GLODER A.,  
 RA SCOTT J.L., GOGHAGEN N.S.M., WEIDMAN J.F., HURST M.A., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii."  
 RL SCIENCE 273:1058-1073(1996).  
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -|- SIMILARITY: STRONG, TO M.JANNASCHII MJ0129 AND MJ0554.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U67507; G1591295; -  
 CC DR TIGR; MJ0587; -  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 17 37 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 118 138 POTENTIAL.  
 SQ SEQUENCE 153 AA; 17733 MW; 88EC711B CRC32;

Query Match 63.0%; Score 34; DB 1; Length 153;  
 Best Local Similarity 63.6%; Pred. No. 11;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
 ||:|:|:|  
 Db 138 IKVLKIKLNLK 148

RESULT 10  
 FLAE\_METJA STANDARD; PRT; 140 AA.  
 ID FLAE\_METJA  
 AC Q58306;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PUTATIVE FLAGELLA-RELATED PROTEIN E.  
 GN FLAE OR MJ0896.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE: 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,

RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.,  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii".  
RL SCIENCE 273:1058-1073(1996).  
CC -1- SIMILARITY: STRONG, TO M.VOLTAE FLAE, ALSO TO FLAD.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL: U67533; G1591572; -  
DR TIGR: MJ0896; -  
DR FLAGELLA.  
KW  
SQ SEQUENCE 140 AA; 15891 MW; 61446FEB CRC32:  
  
Query Match 61.1%; Score 33; DB 1; Length 140;  
Best Local Similarity 70.0%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 KLFKFLKFLK 11  
| | | | | :  
Db 73 KLLKFLAGIK 82  
  
RESULT 11  
RRPB BEV STANDARD; PRT; 2291 AA.  
ID RRPB\_BEV STANDARD; PRT; 2291 AA.  
AC P18458;  
DT 01-NOV-1990 (REL. 16, CREATED)  
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (ORF1B).  
GN POL.  
OS BERNE VIRUS (BEV).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; NIDOVIRALES;  
OC CORONAVIRIDAE; TOROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE P138/72;  
RX MEDLINE; 90356389.  
RA RIJNBAND E.J., DEN BOON J.A., BREDEBEEK P.J., HORZINEK M.C.,  
RA RIJNBAND R., SPAAN W.J.M.;  
RT "The carboxyl-terminal part of the putative Berne virus polymerase is  
RT expressed by ribosomal frameshifting and contains sequence motifs  
RT which indicate that toro- and coronaviruses are evolutionarily  
RT related".  
RL NUCLEIC ACIDS RES. 18:4535-4542(1990).  
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE  
CC + RNA(N).  
CC  
CC -1- THIS PROTEIN IS EXPRESSED BY AN EFFICIENT RIBOSOMAL FRAMESHIFTING  
CC MECHANISM. RIBOSOMAL FRAMESHIFTING IS AN ELEGANT MECHANISM FOR  
CC REGULATING THE SYNTHESIS OF SEVERAL PROTEINS IN A WELL BALANCED  
CC MANNER.  
CC  
CC -1- SIMILARITY: TO CORONAVIRUSES POLYMERASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL: X52374; E23105; -  
DR PIR: S11238; S11238.

KW TRANSFERASE: RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
FT DOMAIN 515 744 POLYMERASE.  
FT NP\_BIND 1099 1374 HELICASE.  
FT NP\_BIND 1101 1108 ATP (BY SIMILARITY).  
FT SIMILAR 845 928 TO ITS HOMOLOG FROM MHV (AA 938-1022)  
FT AND IBV (AA 929-1013).  
FT SIMILAR 1921 2209 TO ITS HOMOLOG FROM MHV (AA 2322-2641)  
FT AND IBV (AA 2274-2598).  
SQ SEQUENCE 2291 AA; 260843 MW; C0283C2D CRC32:  
  
Query Match 61.1%; Score 33; DB 1; Length 2291;  
Best Local Similarity 87.5%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 LKFLKFLK 11  
| | | | |  
Db 523 LKFSKFLK 530  
  
RESULT 12  
VGNE\_CPSMV STANDARD; PRT; 1858 AA.  
ID VGNE\_CPSMV STANDARD; PRT; 1858 AA.  
AC P36312;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE GENOME POLYPROTEIN B (CONTAINS: PROTEASE COFACTOR; MEMBRANE BINDING  
DE PROTEIN; VP; PROTEASE (EC 3.4.22.-); RNA POLYMERASE (EC 2.7.7.48)).  
OS COMPEA SEVERE MOSAIC VIRUS (STRAIN DG) (CPSMV).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; COMOVIRIDAE;  
OC COMOVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93079863.  
RA CHEN X., BRUENING G.;  
RT "Cloned DNA copies of cowpea severe mosaic virus genomic RNAs:  
RT infectious transcripts and complete nucleotide sequence of RNA 1.";  
RL VIROLOGY 191:607-618(1992).  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL: M83830; -; NOT\_ANNOTATED\_CDS.  
DR PIR: A4214; A4214.  
DR PFAM: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR PFAM: PF00910; RNA\_helicase; 1.  
KW POLYPROTEIN; TRANSMEMBRANE; HYDROLASE; PROTEASE; TRANSFERASE;  
KW RNA-DIRECTED RNA POLYMERASE; ATP-BINDING.  
FT CHAIN 1 313 PROTEASE COFACTOR (POTENTIAL).  
FT CHAIN 314 908 MEMBRANE-BINDING PROTEIN (POTENTIAL).  
FT CHAIN 909 936 VPG PROTEIN (POTENTIAL).  
FT CHAIN 937 1146 PROTEASE (POTENTIAL).  
FT CHAIN 1147 1858 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).  
FT TRANSMEM 891 907 POTENTIAL.  
FT NP\_BIND 482 489 ATP (POTENTIAL).  
FT ACT\_SITE 976 976 THIOL PROTEASE (POTENTIAL).  
FT ACT\_SITE 1013 1013 THIOL PROTEASE (POTENTIAL).  
FT ACT\_SITE 1104 1104 THIOL PROTEASE (POTENTIAL).  
SQ SEQUENCE 1858 AA; 208861 MW; 38B826AB CRC32:  
  
Query Match 61.1%; Score 33; DB 1; Length 1858;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
|||||:|:  
Db 1379 KFLKEVFI 1387

## RESULT 13

VIT\_ONCMY STANDARD; PRT: 1659 AA.  
AC Q92093; P79882; Q91190; Q92092;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DE VITELLOGENIN PRECURSOR (VTG) [CONTAINS: LIPOVITELLIN I (LVI);  
DE PHOSVITIN (PV); LIPOVITELLIN II (LVII)].  
GN VTGI.  
OS ONCORHYNCHUS MYKISS (RAINBOW TROUT) (SALMO GAIRDNERI).  
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
CC TELEOSTEI; EUTELEOSTEI; PROTACANTHOPTERYGII; SALMONIFORMES;  
CC SALMONIDAE; ONCORHYNCHUS.  
[1]  
RN RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE; 97017128.  
RA MOUCHEL N., TRICHET V., BETZ A., LE PENNEC J.-P., WOLFF J.;  
RT "Characterization of vitellogenin from rainbow trout (Oncorhynchus  
RT mykiss).";  
RL GENE 174:59-64(1996).  
RN [2]  
RN RP SEQUENCE OF 1208-1659 FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE; 89053022.  
RA LE GUELLEC K., LAWLESS K., VALOTAIRE Y., KRESS M., TENNISWOOD M.;  
RT "vitellogenin gene expression in male rainbow trout (Salmo  
RT gairdneri).";  
RL GEN. COMP. ENDOCRINOL. 71:359-371(1988).  
RN [3]  
RN RP SEQUENCE OF 1331-1603 FROM N.A.  
RX MEDLINE; 96179362.  
RA REN L., LEWIS S.K., LECH J.J.;  
RT "Effects of estrogen and nonylphenol on the post-transcriptional  
RT regulation of vitellogenin gene expression.";  
RL CHEM. BIOL. INTERACT. 100:67-76(1996).  
RN [4]  
RN RP SEQUENCE OF 894-1301 FROM N.A., AND SEQUENCE OF 1089-1093.  
RC TISSUE=LIVER;  
RX MEDLINE; 96326296.  
RA GOUJAS A., TRIPLETT E.L., TABORSKY G.;  
RT "Isolation and characterization of a vitellogenin cDNA from rainbow  
RT trout (Oncorhynchus mykiss) and the complete sequence of a phosvitin  
RT coding segment.";  
RL DNA CELL BIOL. 15:605-616(1996).  
CC -!- FUNCTION: PRECURSOR OF THE MAJOR EGG-YOLK PROTEINS THAT ARE  
CC ORGANISMS.  
CC -!- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD  
CC AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENDOCYTOSIS INTO GROWING  
CC OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE  
CC RESPECTIVE YOLK COMPONENTS LIPOVITELLIN-I, PHOSVITIN,  
CC LIPOVITELLIN-II.  
CC -!- INDUCTION: SYNTHESIZED IN THE LIVER OF OVIPAROUS VERTEBRATES IN  
CC RESPONSE TO STEROID (ESTROGEN) INDUCTION.  
CC -!- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST  
CC HIGHLY PHOSPHORYLATED (10x) PROTEINS IN NATURE.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X92804; E209212; -.

DR EMBL; M27651; G1066855; -.  
DR EMBL; U26703; G1293537; -.  
DR EMBL; S82450; G1703722; -.  
KW GLYCOPROTEIN; PHOSPHORYLATION; STORAGE PROTEIN; SIGNAL.  
FT SIGNAL 1 15  
FT CHAIN 16 1659  
FT CHAIN 16 1659  
FT CHAIN 1089 1145  
FT CHAIN 1146 1659  
FT CHAIN 1089 1089  
FT CARBOHYD 1627 1627  
FT CARBOHYD 1089 1089  
FT DOMAIN 1065 509  
FT DOMAIN 1065 1068  
FT DOMAIN 1094 1112  
FT DOMAIN 1117 1128  
FT VARIANT 1130 1130  
FT CONFLICT 894 894  
FT CONFLICT 1034 1035  
FT CONFLICT 1158 1158  
FT CONFLICT 1208 1218  
FT CONFLICT 1222 1222  
FT CONFLICT 1246 1246  
FT CONFLICT 1254 1254  
FT CONFLICT 1278 1281  
FT CONFLICT 1305 1305  
FT CONFLICT 1322 1322  
FT CONFLICT 1331 1335  
FT CONFLICT 1352 1352  
FT CONFLICT 1419 1419  
FT CONFLICT 1431 1443  
FT CONFLICT 1457 1457  
FT CONFLICT 1470 1470  
FT CONFLICT 1473 1474  
FT CONFLICT 1490 1491  
FT CONFLICT 1491 1491  
FT CONFLICT 1504 1504  
FT CONFLICT 1509 1512  
FT CONFLICT 1546 1546  
FT CONFLICT 1563 1565  
FT CONFLICT 1606 1606  
FT CONFLICT 1622 1622  
SQ SEQUENCE 1659 AA; 183127 MW; 8BC6D0F5 CRC32;  
POTENTIAL.  
VITTELLOGENIN.  
LIPOVITELLIN I.  
PHOSVITIN.  
LIPOVITELLIN II.  
POTENTIAL.  
POTENTIAL.  
POLY-ALA.  
POLY-GLU.  
POLY-SER.  
POLY-SER.  
I -> T.  
F -> L (IN REF. 4).  
SA -> S (IN REF. 4).  
S -> P (IN REF. 4).  
TAYLNKATSR -> LGRPKTSDEP (IN REF. 2).  
M -> T (IN REF. 2).  
T -> N (IN REF. 2).  
E -> G (IN REF. 2).  
RLSW -> L (IN REF. 2).  
Y -> H (IN REF. 2).  
N -> S (IN REF. 2).  
VATSE -> LPHLK (IN REF. 3).  
L -> T (IN REF. 2).  
A -> G (IN REF. 3).  
LLKKDHASEQNHI -> SAEEGSVHLTKTRS (IN  
REF. 2).  
E -> Q (IN REF. 2).  
I -> V (IN REF. 2).  
DN -> EQ (IN REF. 2).  
GK -> KGE (IN REF. 2).  
K -> E (IN REF. 3).  
E -> K (IN REF. 2).  
KYSW -> SNSR (IN REF. 3).  
L -> V (IN REF. 2).  
SCR -> RC (IN REF. 2).  
L -> S (IN REF. 2).  
V -> F (IN REF. 2).

Query Match 61.1%; Score 33; DB 1; Length 1659;  
Best Local Similarity 45.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
:||||:|:  
Db 331 LKFLQFIOLLR 341

## RESULT 14

YG3C\_YEAST STANDARD; PRT: 1122 AA.  
ID YG3C\_YEAST  
AC P53280;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 128.8 KD PROTEIN IN PAS2-PRS5 INTERGENIC REGION.  
GN YG134W  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
CC SACCHAROMYCETACEAE; SACCHAROMYCES.  
[1]  
RN RP SEQUENCE FROM N.A.  
RA VAN DYCK L., SKALA J., DE WERGIFOSSE P., PURNELLE B., TALLA E.,  
RA NAWROCKI A., DEL BINO S., GOFFEAU A.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

Search completed: September 7, 1999, 23:59:18  
Job time: 530 sec

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; 272919; E243526; -  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 1122 AA; 128779 MW; 307CEC10 CRC32;

Query Match 61.1%; Score 33; DB 1; Length 1122;  
Best Local Similarity 70.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
| | | | | | | |  
Db 125 ILLKFLRFL 134

RESULT 15  
YGZB\_YEAST STANDARD; PRT; 558 AA.  
AC P33059;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 64.9 KD PROTEIN IN ADH4 5'REGION (EC 2.4.1.-).  
GN YGL257C OR NR0558.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / FY1679;  
RX MEDLINE; 97127827.  
RA COISSAC E., MAILLIER E., ROBINEAU S., NETTER P.;  
RT "Sequence of a 39,411 bp DNA fragment covering the left end of  
RT chromosome VII of *Saccharomyces cerevisiae*.";  
RL YEAST 12:1555-1562(1996).  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE MN1/YGL257C/YIL014W/YNR059W FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X94357; E215614; -  
DR EMBL; 272779; E243909; -  
KW HYPOTHETICAL PROTEIN; TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN;  
KW TRANSMEMBRANE; SIGNAL-ANCHOR.  
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 7 27 POTENTIAL.  
FT DOMAIN 28 558 LUMENAL (POTENTIAL).  
FT CARBOHYD 187 187 POTENTIAL.  
SQ SEQUENCE 558 AA; 64852 MW; 5CBB32F0 CRC32;

Query Match 61.1%; Score 33; DB 1; Length 558;  
Best Local Similarity 54.5%; Pred. No. 60;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 11  
| | | | | | | |  
Db 190 IRLKVLRLFIK 200

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:47 ; Search time 160.2 Seconds  
(without alignments)  
4.226 Million cell updates/sec

Title: US-09-124-280A-37

Perfect score: 54

Sequence: 1 IKFLKFLKFLK 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	68.5	201	2	O51263 borrelia bu
2	37	68.5	552	8	Q92YM7 rhipicephal
3	36	66.7	185	5	P92014 caenorhabdi
4	36	66.7	951	10	Q43106 phaseolus v
5	36	66.7	953	10	P93265 mesembryant
6	36	66.7	956	10	Q42932 nicotiana p
7	36	66.7	963	10	Q43131 vicia faba
8	36	66.7	956	10	Q43182 solanum tub
9	36	66.7	949	10	Q43243 zea mays (m
10	36	66.7	956	10	Q43001 oryza sativ
11	36	66.7	957	10	Q43002 oryza sativ
12	36	66.7	403	13	O73742 xenopus lae
13	35	64.8	1279	2	O26046 helicobacte
14	35	64.8	290	3	P78815 schizosacch
15	35	64.8	1289	5	Q17174 boophilus m
16	35	64.8	72	5	O23632 caenorhabdi
17	35	64.8	1278	5	O96228 plasmodium
18	34	63.0	260	1	O27551 methanobact
19	34	63.0	132	1	O58773 pyrococcus
20	34	63.0	564	2	O67184 aquifex aeo
21	34	63.0	1119	5	O77374 plasmodium
22	34	63.0	350	5	Q19556 caenorhabdi
23	34	63.0	873	5	O18228 caenorhabdi
24	34	63.0	1687	11	O61204 mus musculu
25	33	61.1	192	2	O51293 borrelia bu
26	33	61.1	286	2	P94922 mycobacteri
27	33	61.1	644	2	O67389 aquifex aeo
28	33	61.1	141	2	O67910 aquifex aeo
29	33	61.1	53	2	O68585 streptococc

30	33	61.1	486	3	O94712 schizosacch
31	33	61.1	45	4	Q14460 homo sapien
32	33	61.1	93	5	O01587 caenorhabdi
33	33	61.1	649	5	O44752 caenorhabdi
34	33	61.1	2657	5	O77380 plasmodium
35	33	61.1	402	5	O45858 caenorhabdi
36	33	61.1	686	10	O40940 chlorarachn
37	33	61.1	4572	10	O40712 oryza sativ
38	33	61.1	72	12	O88417 spiroplasma
39	32.5	60.2	585	2	Q00419 staphylococ
40	32	59.3	129	1	O29640 archaeoglob
41	32	59.3	450	1	O59355 pyrococcus
42	32	59.3	936	2	O25141 helicobacte
43	32	59.3	361	2	O67410 aquifex aeo
44	32	59.3	130	2	O67445 aquifex aeo
45	32	59.3	1850	12	Q9YJU5 bean pod mo

## ALIGNMENTS

RESULT 1					
O51263					
ID	O51263	PRELIMINARY;	PRT;	201 AA.	
AC	O51263;				
DT	01-JUN-1998 (TrEMBLrel. 06, Created)				
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)				
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)				
DE	CONSERVED HYPOTHETICAL PROTEIN.				
GN	BB0247.				
OS	Borrelia burgdorferi (Lyme disease spirochete).				
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ATCC 35210 / B31;				
RX	MEDLINE; 98065943.				
RA	FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,				
RA	LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,				
RA	DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,				
RA	PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,				
RA	VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,				
RA	UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,				
RA	GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,				
RA	SMITH H.O., VENTER J.C.;				
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia				
RT	burgdorferi".				
RL	Nature 390:580-586(1997).				
DR	EMBL; AE001135; AAB91499.1; -.				
DR	TIGR; BB0247; -.				
SQ	SEQUENCE 201 AA; 23138 MW; 1F87F075 CRC32;				

Query Match 68.5%; Score 37; DB 2; Length 201;  
Best Local Similarity 80.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	IKFLKFLKFL 10
DB	185	IAFLKFKFL 194

## RESULT 2

Q92YM7					
ID	Q92YM7	PRELIMINARY;	PRT;	552 AA.	
AC	Q92YM7;				
DT	01-MAY-1999 (TrEMBLrel. 10, Created)				
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)				
DT	01-MAY-1999 (TrEMBLrel. 10, Last annotation update)				
DE	NADH DEHYDROGENASE 5.				
OS	Rhipicephalus sanguineus.				
OG	Mitochondrion.				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.				

```

RN SEQUENCE FROM N.A.
RX MEDLINE; 99083443.
RA BLACK W.C. IV, ROHRDANZ R.L.;
RT "Mitochondrial gene order is not conserved in arthropods: prostriate
and metastriate tick mitochondrial genomes.";
RL Mol. Biol. Evol. 15:1772-1785(1998).
[1]
RP SEQUENCE FROM N.A.
RA BLACK W.C. IV, ROHRDANZ R.L.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF081829; AAD05525.1; -.
KW Mitochondrion.
SQ SEQUENCE 552 AA; 64542 MW; 82C26959 CRC32;

Query Match 68.5%; Score 37; DB 8; Length 552;
Best Local Similarity 54.5%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 11
   :||:|:|:|
Db 464 LKFIFFFLK 474

RESULT 3
P92014 PRELIMINARY; PRT; 185 AA.
AC P92014;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE T02B5.2 PROTEIN.
GN T02B5.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RP SEQUENCE FROM N.A.
RA MCMURRAY A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBDJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-NIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WAYSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z81112; CAB03274.1; -.
SQ SEQUENCE 185 AA; 21299 MW; A87A13B5 CRC32;

Query Match 66.7%; Score 36; DB 5; Length 185;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10
   | | | | |
Db 169 INFLNVLKFL 178

RESULT 4
Q43106 PRELIMINARY; PRT; 951 AA.
ID Q43106

```

```

AC Q43106;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).
GN BHA-1.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
OC Phaseolus.
[1]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA CAMPOS F., MARZ-CASTINEIRA J.R., VILLALBA J.M., SERRANO R.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBDJ databases.
DR EMBL; X85804; CAA59799.1; -.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR MENDEL; 13288; PHAVU; 494; 1.
DR PFAM; PF00122; E1-E2_ATPase; 1.
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.
FT MOD_RES 331 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 951 AA; 104442 MW; 3B569E8C CRC32;

Query Match 66.7%; Score 36; DB 10; Length 951;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10
   | | | | | |
Db 59 KFLKFLGFM 67

RESULT 5
P93265 PRELIMINARY; PRT; 953 AA.
AC P93265;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).
GN PMA.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Caryophyllales; Caryophyllaceae;
OC Mesembryanthemum.
[1]
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA MICHALOWSKI C.B., QUIGLEY-LANDREAU F., BOHNERT H.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBDJ databases.
DR EMBL; U84891; AB41898.1; -.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR MENDEL; 13177; MESCR; 494; 1.
DR PFAM; PF00122; E1-E2_ATPase; 1.
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.
FT MOD_RES 335 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 953 AA; 104924 MW; B11E9AA4 CRC32;

Query Match 66.7%; Score 36; DB 10; Length 953;
Best Local Similarity 77.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10
   | | | | | |
Db 63 KFLKFLGFM 71

RESULT 6
Q42932 PRELIMINARY; PRT; 956 AA.
ID Q42932
AC Q42932;

```

DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE.  
 OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Nicotiana.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89334860.  
 RA BOUTRY M., MICHELET B., GOFFEAU A.;  
 RT "Molecular cloning of a family of plant genes encoding a protein  
 RT homologous to plasma membrane H<sup>+</sup>-translocating ATPases.";  
 RL Biochem. Biophys. Res. Commun. 162:567-574(1989).  
 DR EMBL; M27888; AAA34052.1; -.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 16027; NICP1; 494; mni16027.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105047 MW; FF2D51BE CRC32;

Query Match 66.7%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 Db 61 KFLKFLGFM 69  
 ||||| I:

RESULT 7  
 Q43131 ID Q43131 PRELIMINARY; PRT; 963 AA.  
 AC Q43131;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H(+)-ATPASE.  
 OS Vicia faba (Broad bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Vicia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96012934.  
 RA NAKAJIMA N., SAJI H., AONO M., KONDO N.;  
 RT "Isolation of cDNA for a plasma membrane H(+)-ATPase from guard cells  
 RT of Vicia faba L.";  
 RL Plant Cell Physiol. 36:919-924(1995).  
 DR EMBL; S79323; AAB35314.1; -.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13431; VICfa; 494; 3.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 338 338 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 963 AA; 105965 MW; 2C2C9793 CRC32;

Query Match 66.7%; Score 36; DB 10; Length 963;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 Db 66 KFLKFLGFM 74  
 ||||| I:

RESULT 8

Q43182 ID Q43182 PRELIMINARY; PRT; 956 AA.  
 AC Q43182;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
 GN PHAL.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Solanum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-LEAF.  
 RA HARMS K., WOHNER R.V., SCHULZ B., FROMMER W.B.;  
 RT "Isolation and characterization of P-type H(+)-ATPase genes from  
 RT potato.";  
 RL Plant Mol. Biol. 26:979-988(1994).  
 DR EMBL; X76536; CAA54046.1; -.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 530; SOLTu; 494; 1.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105125 MW; 80AD1D45 CRC32;

Query Match 66.7%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 Db 61 KFLKFLGFM 69  
 ||||| I:

RESULT 9  
 Q43243 ID Q43243 PRELIMINARY; PRT; 949 AA.  
 AC Q43243;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMELrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE.  
 GN MHAI.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Zea.  
 RN [1]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RC STRAIN=D3L;  
 RX MEDLINE; 95003707.  
 RA JIN Y., BENNETZEN J.L.;  
 RT "Integration and nonrandom mutation of a plasma membrane proton  
 RT ATPase gene fragment within the Bsl retroelement of maize.";  
 RL Plant Cell 6:1177-1186(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D3L;  
 RA JIN Y.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U09989; AAB60276.1; -.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 533; ZEAm; 494; 2.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW Hydrolyase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 336 336 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 949 AA; 103900 MW; DD1AA0CF CRC32;

Query Match 66.7%; Score 36; DB 10; Length 949;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 ||||| |  
 Db 64 KFLKFLGFM 72

## RESULT 10

ID Q43001 PRELIMINARY; PRT; 956 AA.  
 AC Q43001;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE H-ATPASE (EC 3.6.1.3) (ADENOSINETRIPHOSPHATASE)  
 DE (ADENYLTRIPHOSPHATASE) (ATP MONOPHOSPHATASE) (TRIPHOSPHATASE).  
 GN OSAL.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eukaryotes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NIPPONBARE K-1;  
 RA WADA M., TAKANO M., KASAMO K.;  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + PHOSPHATE.  
 DR EMBL; D10207; BAA01058.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13241; Oryza; 494.1.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333  
 FT PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105079 MW; 1A9E387B CRC32;

Query Match 66.7%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 ||||| |  
 Db 61 KFLKFLGFM 69

## RESULT 11

ID Q43002 PRELIMINARY; PRT; 957 AA.  
 AC Q43002;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H<sup>+</sup>-ATPASE.  
 GN OSA2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eukaryotes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RX MEDLINE; 95211383.  
 RA OKURA T., WADA M., SAKAKIBARA Y., JEONG K.H., MARUTA I., KAWAMURA Y.,  
 RA KASAMO K.;  
 RT Identification and characterization of a family of genes for the  
 RT plasma membrane H<sup>+</sup>-ATPase of Oryza sativa L.;  
 RT Plant Cell Physiol. 35:1251-1256(1994).  
 DR EMBL; D31843; BAA06629.1; -;  
 DR MENDEL; 13242; Oryza; 494.2.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.

SQ SEQUENCE 957 AA; 105607 MW; 6673B764 CRC32;

Query Match 66.7%; Score 36; DB 10; Length 957;  
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 ||||| |  
 Db 61 KFLKFLGFM 69

## RESULT 12

ID 073742 PRELIMINARY; PRT; 403 AA.  
 AC 073742;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE XOE-2 (FRAGMENT).  
 GN XOE-2.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;  
 OC Xenopus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEAD;  
 RA BALDESSARI D., CONSULEZ G.G.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF040994; AAC15659.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 403 AA; 43800 MW; BB99FC22 CRC32;

Query Match 66.7%; Score 36; DB 13; Length 403;  
 Best Local Similarity 77.8%; Pred. No. 79;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 3 FLKFLKFLK 11  
 ||| ||| |  
 Db 376 FLKILKFKV 384

## RESULT 13

ID 026046 PRELIMINARY; PRT; 1279 AA.  
 AC 026046;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE TYPE IIS RESTRICTION ENZYME R AND M PROTEIN (EC0571R).  
 GN HP1517.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695;  
 RX MEDLINE; 97394467.  
 RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
 RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
 RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKLEY E.K.,  
 RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATHEY L., WALLIN E.,  
 RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,  
 RA VENTER J.C.;  
 RT The complete genome sequence of the gastric pathogen Helicobacter  
 RT pylori.;  
 RT Nature 388:539-547(1997).  
 RL EMBL; A5000649; AAD08556.1; -.

DR TIGR: HP1517; -  
KW Hypothetical protein.  
SQ SEQUENCE 1279 AA; 149714 MW; 074DBDA7 CRC32;

Query Match 64.8% Score 35; DB 2; Length 1279;  
Best Local Similarity 80.0%; Pred. No. 3e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 11  
:|||||  
DB 738 KELKFKKFLK 747

## RESULT 14

P78815  
ID P78815 PRELIMINARY; PRT; 290 AA.  
AC P78815;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE FISSION YEAST (FRAGMENT).  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PR745;  
RX MEDLINE; 98162722.  
RA YOSHIOKA S., KATO K., NAKAI K., OKAYAMA H., NOJIMA H.;  
RT "Identification of open reading frames in Schizosaccharomycetes pombe  
CDNAS.";  
RL DNA Res. 4:363-369(1997).  
DR EMBL: D89164; BAA13826.1; -.  
DR PFAM: PF00018; SH3; 2.  
FT NON\_TER 1  
SQ SEQUENCE 290 AA; 33434 MW; 86761774 CRC32;

Query Match 64.8% Score 35; DB 3; Length 290;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 FLKFLKFLK 11  
:|||||  
DB 105 YLFLQLFLK 113

## RESULT 15

Q17174  
ID Q17174 PRELIMINARY; PRT; 1289 AA.  
AC Q17174;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE GP80 (FRAGMENT).  
OS Boophilus microplus (Cattle tick).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VEERONGPILLY;  
RA IRVING D., SMITH D.R.J., SHARP P., RIDING G., BRISCOE S.,  
RA TELLAM R.L.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U49934; AAA2143.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 1289 AA; 147185 MW; A82B0E07 CRC32;

Query Match 64.8% Score 35; DB 5; Length 1289;  
Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 IKFLKFLKF 9  
:|||||  
DB 889 LKFLSYLKF 897  
Search completed: September 7, 1999, 22:47:48  
Job time: 7968 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:07 ; Search time 135.78 Seconds  
(without alignments)  
1.221 Million cell updates/sec

Title: US-09-124-280A-38

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	32	1 R33514	Generic peptide fo
2	35	100.0	11	1 R33532	Peptide for treati
3	35	100.0	7	1 R33533	Peptide for treati
4	35	100.0	11	1 R39294	Endotoxin lipid A
5	35	100.0	7	1 R39295	Endotoxin lipid A
6	35	100.0	7	1 R71178	Peptide neutralisi
7	35	100.0	10	1 R71179	Peptide neutralisi
8	35	100.0	12	1 R71791	Peptide neutralisi
9	35	100.0	11	1 W21625	Antibiotic potenti
10	35	100.0	7	1 W21626	Antibiotic potenti
11	35	100.0	10	1 W21629	Antibiotic potenti
12	35	100.0	7	1 W21595	Antibiotic potenti
13	35	100.0	10	1 W21596	Antibiotic potenti
14	35	100.0	12	1 W21608	Antibiotic potenti
15	30	85.7	6	1 R33515	Peptide for treati
16	29	82.9	1279	1 W98761	H. pylori GPO 109
17	28	80.0	116	1 R65492	Pyruvate dehydrog
18	28	80.0	116	1 W52989	Homo sapiens clone
19	27	77.1	405	1 R05713	Arginine deaminase
20	27	77.1	404	1 R48674	Putative peach pol
21	27	77.1	869	1 R53732	S. cerevisiae plcl
22	27	77.1	10	1 R71781	Peptide neutralisi
23	27	77.1	348	1 R85333	Human prostate/col
24	27	77.1	347	1 R85334	Human prostate/col
25	27	77.1	2165	1 R94930	RSV RNA-dependent
26	27	77.1	420	1 W10168	Human regulator of
27	27	77.1	9	1 W21628	Antibiotic potenti
28	27	77.1	10	1 W21598	Antibiotic potenti
29	27	77.1	348	1 W15093	Form 1 of prostate
30	27	77.1	347	1 W15094	Form 2 of prostate
31	27	77.1	1580	1 W57412	Homo sapiens sulph
32	27	77.1	2166	1 W48715	RSV isolate 18537
33	27	77.1	2166	1 W48716	RSV vaccine 2833F
34	27	77.1	2166	1 W48717	RSV vaccine 2833F
35	27	77.1	2166	1 W48718	RSV vaccine 2833F
36	27	77.1	2166	1 W48719	RSV revertant 2833
37	27	77.1	2166	1 W48714	RSV revertant 2820
38	27	77.1	2165	1 W77015	RSV isolate 2B will
39	27	77.1	2165	1 W77015	Respiratory syncyt
40	27	77.1	500	1 W83330	Thermus flavus amy
41	27	77.1	107	1 Y11315	S. pneumoniae prot
42	27	77.1	333	1 Y11025	H. pylori ORF 029e
43	27	77.1	1522	1 Y03183	Topoisomerase II b
			2166	1 Y02005	L protein of RSV w

ALIGNMENTS

RESULT 1

R33514  
ID R33514 standard; peptide; 32 AA.  
AC R33514;  
DT 07-JUL-1993 (first entry)  
DE Generic peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PS Claim 2; Page 32; 39pp; English.  
CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
CC where n = 1-10, the peptide is useful for treating or preventing  
CC septic shock, mixing with polymyxin B to reduce its toxicity;  
CC removing endotoxins from blood, sera or other fluids (in vivo  
CC or in vitro); controlling release of cytokines induced by  
CC endotoxins; as diagnostic reagents to detect and quantify toxins  
CC in blood or sera; preparing non-toxic antigenic complexes of  
CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
CC bacterial meningitis and HIV-related infections. The usual dose is  
CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 32 AA;

Query Match 100.0%; Score 35; DB 1; Length 32;  
Best Local Similarity 100.0%; Pred. No. 0.66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

DB 2 KFLKFLK 8

RESULT 2

R33532  
ID R33532 standard; peptide; 11 AA.  
AC R33532;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M.  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated  
 PS Claim 10; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 Db 2 KFLKFLK 8

RESULT 3  
 R33533  
 ID R33533 standard; peptide; 7 AA.  
 AC R33533;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 11; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 Db 1 KFLKFLK 7

RESULT 4  
 R39294  
 ID R39294 standard; peptide; 11 AA.  
 AC R39294;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 Db 2 KFLKFLK 8

RESULT 5  
 R39295  
 ID R39295 standard; peptide; 7 AA.  
 AC R39295;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent



PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 11: Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products. 7 AA;  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 1 KFLKFLK 7  
 |||||

RESULT 6  
 R71778  
 ID R71778 standard; peptide; 7 AA.  
 AC R71778;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 28-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 9: Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 10 AA;  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 1 KFLKFLK 7  
 |||||

RESULT 8  
 R71791  
 ID R71791 standard; peptide; 12 AA.  
 AC R71791;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..12  
 PD PN WO9503327-A.  
 PF 02-FEB-1995.  
 PR 21-JUL-1994; E02413.  
 PR 28-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 22; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 12 AA;  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 1 KFLKFLK 7  
 |||||

RESULT 7  
 R71779  
 ID R71779 standard; peptide; 10 AA.  
 AC R71779;  
 DT 01-OCT-1995 (first entry)

DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 28-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 10 AA;  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 1 KFLKFLK 7  
 |||||

RESULT 8  
 R71791  
 ID R71791 standard; peptide; 12 AA.  
 AC R71791;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..12  
 PD PN WO9503327-A.  
 PF 02-FEB-1995.  
 PR 21-JUL-1994; E02413.  
 PR 28-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 22; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)n, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides. 12 AA;  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 |||||  
 Db 2 KFLKFLK 8

## RESULT 9

W21625 ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 35; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 |||||  
 Db 2 KFLKFLK 8

## RESULT 10

W21626 ID W21626 standard; peptide; 7 AA.  
 AC W21626;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #38.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 42; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 |||||  
 Db 1 KFLKFLK 7

## RESULT 11

W21629 ID W21629 standard; peptide; 10 AA.  
 AC W21629;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #41.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
 |||||  
 Db 2 KFLKFLK 8

## RESULT 12

W21595 ID W21595 standard; peptide; 7 AA.  
 AC W21595;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #7.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 11; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 | | | | |  
 DB 1 KFLKFLK 7

RESULT 13  
 W21596  
 ID W21596 standard; peptide; 10 AA.  
 AC W21596;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #8.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996; E02313.  
 PF 29-MAY-1996; US-456112.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 12; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 35; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 | | | | |  
 DB 1 KFLKFLK 7

RESULT 14  
 W21608  
 ID W21608 standard; peptide; 12 AA.  
 AC W21608;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #20.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT disulfide\_bond 1. .12  
 PN W09638163-A1.  
 PD 05-DEC-1996;  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 24; Page 26; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 35; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 | | | | |  
 DB 2 KFLKFLK 8

RESULT 15  
 R33515  
 ID R33515 standard; peptide; 6 AA.  
 AC R33515;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B; Group I.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT region 1. .3  
 FT /note- "repeat region"  
 FT region 4. .6  
 FT /note- "repeat region"  
 FT ZAG200943-A.  
 PD 25-NOV-1992;  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Example; Page 5; 39pp; English.  
 CC The (Group I) peptide is an example of a generic peptide of formula  
 CC R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where n = 1-100  
 CC and each R is H, an amino acid residue or a fatty acid residue.  
 CC The peptide is useful for treating or preventing septic shock,  
 CC mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 6 AA;

Query Match 85.7%; Score 30; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKFL 6  
|||  
Db 1 KFLKFL 6

Search completed: September 7, 1999, 22:50:08  
Job time: 7889 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:50 ; Search time 82.37 Seconds  
(without alignments)  
0.839 Million cell updates/sec

Title: US-09-124-280A-38

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*

4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	11	1	US-08-049-871-8
2	35	100.0	11	1	Sequence 8, Appl
3	35	100.0	11	1	Sequence 9, Appl
4	35	100.0	11	1	Sequence 8, Appl
5	35	100.0	7	1	US-08-280-397-8
6	35	100.0	7	1	US-08-280-397-9
7	35	100.0	7	1	US-08-097-830E-7
8	35	100.0	10	1	US-08-097-830E-8
9	35	100.0	12	1	US-08-097-830E-20
10	35	100.0	7	2	US-08-456-112B-7
11	35	100.0	10	2	US-08-456-112B-8
12	35	100.0	12	2	US-08-456-112B-20
13	35	100.0	11	2	US-08-456-112B-37
14	35	100.0	7	2	US-08-456-112B-38
15	35	100.0	10	2	US-08-456-112B-41
16	28	80.0	887	1	US-08-215-709-1
17	27	77.1	116	1	US-08-702-344-21
18	27	77.1	10	1	US-08-097-830E-10
19	27	77.1	2165	2	US-08-514-975B-2
20	27	77.1	10	2	US-08-456-112B-10
21	27	77.1	9	2	US-08-456-112B-40
22	27	77.1	2165	3	PCT-US95-12507-2
23	26	74.3	13	1	US-08-465-325-135
24	26	74.3	773	2	US-08-524-757-6
25	26	74.3	772	2	US-08-524-757-12
26	26	74.3	185	2	US-08-338-543-2
27	26	74.3	331	2	US-08-878-989-21
28	25	71.4	21	1	US-07-908-455A-63
29	25	71.4	287	1	US-08-457-245-3
30	25	71.4	21	1	US-08-434-120-89
31	25	71.4	21	1	US-08-465-325-87
32	25	71.4	10	1	US-08-465-325-121
33	25	71.4	11	1	US-08-465-325-122
34	25	71.4	14	1	US-08-465-325-125
35	25	71.4	527	2	US-08-246-583-2
36	25	71.4	737	2	US-08-185-432-2
37	25	71.4	737	2	US-08-185-432-4
38	25	71.4	224	2	US-08-185-432-7
39	25	71.4	7	2	US-08-944-133-9

40	25	71.4	14	2	US-08-944-133-10	Sequence 10, Appl
41	25	71.4	21	2	US-08-944-133-11	Sequence 11, Appl
42	25	71.4	28	2	US-08-944-133-12	Sequence 12, Appl
43	25	71.4	14	2	US-08-944-133-45	Sequence 45, Appl
44	25	71.4	21	2	US-08-944-133-46	Sequence 46, Appl
45	25	71.4	28	2	US-08-944-133-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1  
US-08-049-871-8  
; Sequence 8, Application US/08049871  
; Patent No. 5358933  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA: US/08/049,871  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/658,744  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-049-871-8

Query Match 100.0%; Score 35; DB 1;  
Best Local Similarity 100.0%; Pred. No. 0.38; Length 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 2  
US-07-819-893-8  
; Sequence 8, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
;; US-07-819-893-8

Query Match 100.0%; Score 35; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 3  
US-07-819-893-9  
;; Sequence 9, Application US/07819893  
;; Patent No. 5371186  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; US-07-819-893-9

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 4  
US-08-280-397-8  
;; Sequence 8, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-280-397-8

Query Match 100.0%; Score 35; DB 1; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 5  
US-08-280-397-9  
; Sequence 9, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280,397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-280-397-9

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 6  
US-08-097-830E-7  
; Sequence 7, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-7

Query Match 100.0%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 7  
US-08-097-830E-8  
; Sequence 8, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: POIRO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-8

Query Match 100.0%; Score 35; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 8  
US-08-097-830E-20  
Sequence 20, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:

APPLICANT: Poirio, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity Of Lipid A  
NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097.830E

FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:

OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-20

Query Match 100.0%; Score 35; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 9  
US-08-456-112B-7  
Sequence 7, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Poirio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-7

Query Match 100.0%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 10  
US-08-456-112B-8  
Sequence 8, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:

APPLICANT: Poirio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-8

Query Match 100.0%; Score 35; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
DB 1 KFLKFLK 7

RESULT 11  
US-08-456-112B-20  
Sequence 20, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-20

Query Match 100.0%; Score 35; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7  
|||||  
DB 2 KFLKFLK 8  
RESULT 12  
US-08-456-112B-37  
Sequence 37, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-37

Query Match 100.0%; Score 35; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
DB 2 KFLKFLK 8

RESULT 13  
US-08-456-112B-38  
Sequence 38, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POIRO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

Wed Sep 8 09:44:44 1999

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-38

Query Match 100.0%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 1 KFLKFLK 7

RESULT 14  
US-08-456-112B-41  
Sequence 41, Application US/08/456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-41

Query Match 100.0%; Score 35; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 2 KFLKFLK 8

RESULT 15  
US-08-215-709-1  
Sequence 1, Application US/08215709  
Patent No. 5432071  
GENERAL INFORMATION:  
APPLICANT: ICHIKAWA, Toshio  
APPLICANT: KOYAMA, Yasuji  
APPLICANT: OTAKE, Hideko  
APPLICANT: NAKANO, Eiichi  
TITLE OF INVENTION: Variant E1 Protein Gene For Pyruvate  
TITLE OF INVENTION: Dehydrogenase Complex And Variant E1 Protein Of Pyruvate  
TITLE OF INVENTION: Dehydrogenase Complex  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/215,709  
FILING DATE: 22-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jean B. Fordis  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04853.0011-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 887 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-215-709-1

Query Match 80.0%; Score 28; DB 1; Length 887;  
Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
Db 207 KFLKYLE 213

Search completed: September 7, 1999, 23:07:51

Job time: 1749 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:55 ; Search time 105.14 Seconds  
(without alignments)  
2.667 Million cell updates/sec

Title: US-09-124-280A-38

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	91.4	236	2 A56509	peroxisomal membra
2	31	88.6	236	2 C71148	hypothetical prote
3	30	85.7	956	2 A45506	H+-transporting Ar
4	30	85.7	956	2 A43637	H+-transporting Ar
5	30	85.7	956	2 S50751	H+-transporting Ar
6	30	85.7	951	2 S52728	H+-transporting Ar
7	30	85.7	318	2 D64417	transformation sen
8	30	85.7	260	2 A69068	2-hydroxyhepta-2,4
9	30	85.7	141	2 C71702	hypothetical prote
10	30	85.7	949	2 T02083	H+-transporting Ar
11	30	85.7	394	2 S64505	HGH1 protein - yea
12	30	85.7	309	2 S61665	probable membrane
13	30	85.7	1817	2 D71606	hypothetical prote
14	30	85.7	193	2 B70417	conserved hypothet
15	30	85.7	141	2 B70485	hypothetical prote
16	29	82.9	370	2 H70423	oxygen-independent
17	29	82.9	564	2 S50520	hexose transport p
18	29	82.9	564	2 S63405	hexose transport p
19	29	82.9	312	2 G71146	hypothetical prote
20	29	82.9	569	2 A45624	trophozoite cystel
21	29	82.9	2291	2 S11238	polymerase - Berne
22	29	82.9	1279	2 E64709	type IIS restricti
23	29	82.9	121	2 C71696	hypothetical prote
24	29	82.9	475	2 H71660	nitrogen assimilat
25	29	82.9	305	2 S77033	hypothetical prote
26	29	82.9	286	2 B70614	probable mnaAl pro
27	29	82.9	287	2 B64201	hypothetical prote
28	29	82.9	346	2 I64237	hypothetical prote
29	29	82.9	140	2 H64411	hypothetical prote
30	29	82.9	337	2 S53846	ribosomal protein
31	28	80.0	987	1 DESCPV	pyruvate dehydroge
32	28	80.0	201	2 S62757	NADH dehydrogenase
33	28	80.0	718	2 C73088	glycogen operon pr
34	28	80.0	360	2 C70185	carotenoid biosynt
35	28	80.0	448	2 B69745	phosphoglucomutase
36	28	80.0	235	2 B70009	two-component resp
37	28	80.0	317	2 D64233	hypothetical prote
38	28	80.0	558	2 S8614	hypothetical prote
39	28	80.0	2167	2 S50658	bud emergence prot

#### ALIGNMENTS

RESULT 1

A56509

peroxisomal membrane protein PMP27 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein AOE236; protein O0454; protein YOL147c

C:Species: Saccharomyces cerevisiae

C>Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 05-Dec-1997

C:Accession: A56509; B56509; S60389; S62134; S62148; S66844; S57677

R:Erdmann, R.; Blobel, G.

J. Cell Biol. 128, 509-523, 1995

A:Title: Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae lacking the

A:Reference number: A56509; MUID:95164555

A:Accession: A56509

A:Molecule type: DNA

A:Residues: 1-236 <ERD>

A:Cross-references: GB:X81465; NID:9683700; PID:9683701

A:Accession: B56509

A:Molecule type: protein

A:Residues: 2-28 <ER2>

R:Casamayor, A.; Aldea, M.; Casas, C.; Herrero, E.; Gamo, F.J.; Lafuente, M.J.; Gance

Yeast 11, 1281-1288, 1995

A:Title: DNA sequence analysis of a 13 kbp fragment of the left arm of yeast chromoso

A:Reference number: S60385; MUID:96132030

A:Accession: S60389

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-236 <CAW>

A:Cross-references: EMBL:248239; NID:g1163073; PID:g886949

R:Marshall, P.A.; Krimkevich, Y.I.; Lark, R.H.; Dyer, J.M.; Veenhuis, M.; Goodman, J.

J. Cell Biol. 129, 345-355, 1995

A:Title: Pmp27 promotes peroxisomal proliferation.

A:Reference number: S62134

A:Accession: S62134

A:Molecule type: DNA

A:Residues: 1-236 <MAR>

A:Cross-references: EMBL:246846; NID:g791131; PID:g791132

A:Accession: S62148

A:Molecule type: protein

A:Residues: 2-30;192-207 <MAW>

R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66814

A:Accession: S66844

A:Molecule type: DNA

A:Residues: 1-236 <ARI>

A:Cross-references: EMBL:274889; NID:g1420049; PID:e251925; PID:g1420050; MIPS:YOL147

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:PEX11; PMP27

A:Cross-references: MIPS:YOL147c; SGD:S0003507

A:Map position: 15L

C:Function:

A:Description: promotion of peroxisomal proliferation

C:Superfamily: peroxisomal membrane protein PMP27

C:Keywords: peroxisome; transmembrane protein

F:2-236/Product: peroxisomal membrane protein PMP27 #status experimental <MAT>

Query Match 91.4%; Score 32; DB 2; Length 236;

Best Local Similarity 85.7%; Pred. No. 12;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

|||||  
Db 60 KFLFLK 66

RESULT 2  
C71148  
hypothetical protein PH0397 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 14-Aug-1998  
C:Accession: C71148  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 53-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137  
A:Accession: C71148  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-395 <KAW>  
A:Cross-references: CB:AP000002; NID:g3236129; PID:d1030415; PID:g3256789  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH0397

Query Match 88.68; Score 31; DB 2; Length 395;  
Best Local Similarity 85.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
|||||  
Db 237 KLEFLK 243

RESULT 3  
A45506  
H+-transporting ATPase (EC 3.6.1.35) LHAI - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998  
C:Accession: A45506  
R:Ewing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.  
Plant Physiol. 94, 1874-1881, 1990  
A:Title: Molecular cloning of tomato plasma membrane H+-ATPase.  
A:Reference number: A45506  
A:Accession: A45506  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-956 <EWI>  
A:Cross-references: GB:M60166; NID:q170463; PID:q170464  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||  
Db 61 KFLKFL 66

RESULT 4  
A43637  
H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco  
C:Species: Nicotiana glauca (curled-leaved tobacco)  
C>Date: 11-Feb-1993 #sequence\_revision 12-Mar-1993 #text\_change 17-Mar-1999  
C:Accession: A43637  
R:Boutry, M.; Michelet, B.; Goffeau, A.

Biochem. Biophys. Res. Commun. 162, 567-574, 1989  
A:Title: Molecular cloning of a family of plant genes encoding a protein homologous to  
A:Reference number: A43637; MUID:89334860  
A:Accession: A43637  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-956 <BOU>  
A:Cross-references: GB:M27888; NID:q170205; PID:q170206  
A:Note: the authors translated the codon CGG for residue 131 as Trp  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||  
Db 61 KFLKFL 66

RESULT 5  
S50751  
H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato  
C:Species: Solanum tuberosum (potato)  
C>Date: 14-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 24-Apr-1998  
C:Accession: S50751; S39313  
R:Harms, K.; Woehner, R.V.; Schulz, B.; Frommer, W.B.  
Plant Mol. Biol. 26, 979-988, 1994  
A:Title: Isolation and characterization of P-type H(+)-ATPase genes from potato.  
A:Reference number: S50751  
A:Accession: S50751  
A:Molecule type: mRNA  
A:Residues: 1-956 <HAR>  
A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003  
A:Note: the nucleotide is not given  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>  
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:427/Binding site: ATP (lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 956;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||  
Db 61 KFLKFL 66

RESULT 6  
S52728  
H+-transporting ATPase (EC 3.6.1.35) - kidney bean  
C:Species: Phaseolus vulgaris (kidney bean)  
C>Date: 19-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 24-Apr-1998  
C:Accession: S52728  
R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S52728  
A:Accession: S52728  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-951 <CAM>  
A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250  
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain  
C:Keywords: ATP; hydrolase; phosphoprotein  
F:483-655/Domain: ATPase nucleotide-binding domain homology <ATN>

F:331/Active site: Asp (aspartylphosphate intermediate) #status predicted  
F:425/Binding site: ATP (Lys) #status predicted

Query Match 85.7%; Score 30; DB 2; Length 951;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
||| |||  
Db 59 KFLKFL 64

RESULT 7  
D64417  
transformation sensitive protein homolog - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Sep-1998  
C:Accession: D64417  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: D64417  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <BUL>  
A:Cross-references: GB:U67537; GB:L77117; NID:gl591605; PID:gl591608; TIGR:MJ0940; PID:9  
C:Genetics:  
A:Map position: FOR869481-870437  
C:Superfamily: transformation sensitive protein; tetratricopeptide repeat homology  
F:17-50/Domain: tetratricopeptide repeat homology <Tt1>  
F:51-83/Domain: tetratricopeptide repeat homology #status atypical <Tt2>  
F:84-117/Domain: tetratricopeptide repeat homology <Tt3>  
F:118-151/Domain: tetratricopeptide repeat homology <Tt4>  
F:152-185/Domain: tetratricopeptide repeat homology <Tt5>  
F:186-219/Domain: tetratricopeptide repeat homology <Tt6>  
F:255-288/Domain: tetratricopeptide repeat homology <Tt7>

Query Match 85.7%; Score 30; DB 2; Length 318;  
Best Local Similarity 85.7%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFL 7  
||| |||  
Db 51 KFLKFL 57

RESULT 8  
A69068  
2-hydroxyhepta-2,4-diene-1,7-dioate isomerase - Methanobacterium thermoautotrophicum (str  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 13-Sep-1998  
C:Accession: A69068  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
; Glu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanli, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A:Reference number: A69000; MUID:98037514  
A:Accession: A69068  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-260 <MTH>  
A:Cross-references: GB:AE000911; GB:AE000666; NID:g2622623; PID:g2622625  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1507  
A:Start codon: GTG

C:Superfamily: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Query Match 85.7%; Score 30; DB 2; Length 260;  
Best Local Similarity 71.4%; Pred. No. 34;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 7  
||| |||  
Db 5 KFLKFL 11

RESULT 9  
C71702  
hypothetical protein RP436 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 21-Nov-1998  
C:Accession: C71702  
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630  
A:Accession: C71702  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-141 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3860788; PID:el342737; PID:g386099  
A:Experimental source: strain Madrid E  
C:Genetics:  
A:Gene: RP436

Query Match 85.7%; Score 30; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
||| |||  
Db 118 KFLKFL 123

RESULT 10  
T02083  
H+-transporting ATPase (EC 3.6.1.35) Mha1 - maize  
C:Species: Zea mays (maize)  
C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 26-Feb-1999  
C:Accession: T02083  
R:Jin, Y.; Bennetzen, J.L.  
Plant Cell 6, 1177-1186, 1994  
A:Title: Integration and nonrandom mutation of a plasma membrane proton ATPase gene f  
A:Reference number: Z14555  
A:Accession: T02083  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-949 <JIN>  
A:Cross-references: EMBL:U09989; NID:g507770; PID:g533775  
A:Experimental source: strain D3L  
C:Genetics:  
A:Gene: Mha1  
A:Map position: 2L  
A:Introns: 21/3; 61/3; 94/3; 140/3; 185/3; 245/3; 314/3; 354/3; 394/3; 435/3; 470/3;  
C:Function:  
A:Description: transports protons across the plasma membrane to regulate intracellular  
C:Keywords: ATP; hydrogen ion transport; hydrolase; phosphoprotein

Query Match 85.7%; Score 30; DB 2; Length 949;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
||| |||  
Db 64 KFLKFL 69

```

RESULT 11
S64505
HGHI protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G7538; protein YGR187c
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 06-Feb-1998
A:Accession: S64505; S59348
R:Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64499
A:Accession: S64505
A:Molecule type: DNA
A:Residues: 1-394 <ARR>
A:Cross-references: EMBL:Z72972; NID:gl323331; PID:e243565; PID:gl323332; MIPS:YGR187c
A:Experimental source: Strain S288C
R:Sun, Z.; Liang, J.; Hampsey, M.
submitted to the EMBL Data Library, May 1995
A:Description: The HGHI gene encodes a protein with structural similarity to mammalian H
A:Reference number: S59348
A:Accession: S59348
A:Molecule type: DNA
A:Residues: 111-394 <SUN>
A:Cross-references: EMBL:U27358; NID:gl002515; PID:gl002516
C:Genetics:
A:Gene: SGD:HGHI
A:Cross-references: SGD:S0003419; MIPS:YGR187c
A:Map position: 7R
C:Keywords: transmembrane protein
F:296-312/Domain: transmembrane #status predicted <TM>

Query Match 85.7%; Score 30; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6
DB 91 KFLKFL 96

RESULT 12
S61665
probable membrane protein YOR107w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O3224; hypothetical protein YOR3224w
C:Species: Saccharomyces cerevisiae
C:Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 21-Nov-1997
A:Accession: S61665; S66992
R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Sander, C.; Valencia
submitted to the EMBL Data Library, December 1995
A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
A:Reference number: S61643
A:Accession: S61665
A:Molecule type: DNA
A:Residues: 1-309 <BEN>
A:Cross-references: EMBL:X94335; NID:gl262139; PID:e217829; PID:gl1164952
R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoerq, W.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66965
A:Accession: S66992
A:Molecule type: DNA
A:Residues: 1-309 <VOS>
A:Cross-references: EMBL:Z75015; NID:gl420290; PID:e252009; PID:gl420291; MIPS:YOR107w
A:Experimental source: Strain S288C
C:Genetics:
A:Map position: 15R
C:Keywords: transmembrane protein
F:129-145/Domain: transmembrane #status predicted <TM>

Query Match 85.7%; Score 30; DB 2; Length 309;
Best Local Similarity 85.7%; Pred. No. 41;

```

```

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7
DB 50 KFYKFLK 56

RESULT 13
D71606
hypothetical protein PFB0800c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
A:Accession: D71606
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: D71606
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1817 <GAR>
A:Cross-references: GB:AE001418; GB:AE001362; NID:g3845275; PID:g3845280; TIGR:PFB080
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0800c

Query Match 85.7%; Score 30; DB 2; Length 1817;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7
DB 810 FLKFLK 815

RESULT 14
B70417
conserved hypothetical protein aq_1350 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
A:Accession: B70417
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: the complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: B70417
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <AQF>
A:Cross-references: GB:AE000736; NID:g2983763; PID:g2983778; GB:AE000657
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_1350

Query Match 85.7%; Score 30; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6
DB 188 KFLKFL 193

RESULT 15
B70485
hypothetical protein aq_2157 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
A:Accession: B70485

```



R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: B70485  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-141 <AQF>  
A:Cross-references: GB:AE000776; NID:g2984355; PID:g2984366; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_2157

Query Match 85.7%; Score 30; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFL 6  
| | | | |  
Db 38 KFLKFL 43

Search completed: September 7, 1999, 23:22:57  
Job time: 964 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:18 ; Search time 71.09 Seconds  
(without alignments)  
2.783 Million cell updates/sec

Title: US-09-124-280A-38

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	91.4	235	1	PEXB_YEAST
2	30	85.7	394	1	HGH1_YEAST
3	30	85.7	956	1	PMAL_LYCES
4	30	85.7	103	1	RR6_CYACA
5	30	85.7	318	1	Y940_METJA
6	29	82.9	569	1	CYSP_PLAFA
7	29	82.9	140	1	FLAE_METJA
8	29	82.9	564	1	HXTG_YEAST
9	29	82.9	564	1	HXTG_YEAST
10	29	82.9	2291	1	RRPB_BEV
11	29	82.9	337	1	RT07_ACACA
12	29	82.9	287	1	Y011_MYCGE
13	29	82.9	346	1	Y343_MYCGE
14	28	80.0	2167	1	BEM2_YEAST
15	28	80.0	943	1	DIP2_YEAST
16	28	80.0	233	1	DP3E_BUCAP
17	28	80.0	201	1	NU6M_CYACA
18	28	80.0	578	1	ODP1_ECOLI
19	28	80.0	578	1	PSP2_YEAST
20	28	80.0	599	1	VE1_HPVO4
21	28	80.0	317	1	Y302_MYCGE
22	28	80.0	558	1	YCX9_ASTLO
23	28	80.0	628	1	YEL1_SCHPO
24	28	80.0	823	1	YHX2_YEAST
25	27	77.1	279	1	ATPG_MYCGE
26	27	77.1	355	1	CAPH_STAAU
27	27	77.1	626	1	CC23_YEAST
28	27	77.1	526	1	CPF5_RAT
29	27	77.1	817	1	ENPL_CATRO
30	27	77.1	890	1	IMB2_HUMAN
31	27	77.1	451	1	MURD_BORBU
32	27	77.1	404	1	NANH_CLOSO
33	27	77.1	269	1	NIVA_CLOPA
34	27	77.1	869	1	PLC1_YEAST
35	27	77.1	629	1	PR39_YEAST
36	27	77.1	469	1	RGS7_BOVIN
37	27	77.1	420	1	RGS7_HUMAN
38	27	77.1	469	1	RGS7_MOUSE
39	27	77.1	179	1	RL15_ARCFU
40	27	77.1	178	1	RL5_ARCFU
41	27	77.1	2238	1	RRPL_BUNYW
42	27	77.1	2165	1	RRPL_HRSVA
43	27	77.1	195	1	RUBY_CLOPE

ALIGNMENTS

RESULT 1	PEXB_YEAST	STANDARD;	PRT;	235 AA.
AC	Q12462;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DE	PEROXISOMAL MEMBRANE PROTEIN PMP27 (PEROXIN-11).			
GN	PEX11 OR PMP27 OR PMP24 OR YOL147C			
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).			
OC	EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;			
OC	SACCHAROMYCETACEAE; SACCHAROMYCES.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.			
RC	STRAIN-S288C;			
RX	MEDLINE; 95164555.			
RA	ERDMANN R., BLOBEL G.;			
RT	"Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae			
RT	lacking the peroxisomal membrane protein Pmp27p.";			
RL	J. CELL BIOL. 128:509-523(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-29 AND 191-206.			
RC	STRAIN-GRF88;			
RX	MEDLINE; 95238534.			
RA	MARSHALL P.A., KRIMKEVICH Y.I., LARK R.H., DYER J.M., VEENHUIS M.,			
RA	GOODMAN J.M.;			
RT	"Pmp27 promotes peroxisomal proliferation.";			
RL	J. CELL BIOL. 129:345-353(1995).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-S288C / FY1679;			
RX	MEDLINE; 96132030.			
RA	CASANAYOR A., ALDEA M., CASAS C., HERRERO E., GAMO F.J.,			
RA	LAFUENTE M.J., GANCEDO C., ARINO J.;			
RT	"DNA sequence analysis of a 13 kbp fragment of the left arm of yeast			
RT	chromosome XV containing seven new open reading frames.";			
RL	YEAST 11:1281-1288(1995).			
CC	-1- FUNCTION: INVOLVED IN PEROXISOMAL PROLIFERATION. COULD PARTICIPATE			
CC	IN PEROXISOMAL ELONGATION OR FISSION. MAY BE INVOLVED IN PARCELING			
CC	OF PEROXISOMES INTO REGULAR QUANTA.			
CC	-1- SUBCELLULAR LOCATION: PEROXISOMAL MEMBRANE-ASSOCIATED.			
CC	-1- SIMILARITY: STRONG, TO C.BOIDINII PEX11.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@lsb-sib.ch).			
CC	-----			
CC	EMBL; Z48239; G886949; -			
CC	EMBL; Z74889; E251925; -			
DR	EMBL; X81465; G683701; -			
DR	EMBL; Z46846; G791132; -			
DR	SGD; L0002650; PMP27.			
KW	PEROXISOME; MEMBRANE.			
FT	INIT_MET 0			
SEQUENCE	235 AA; 26744 MW; 228BDA38 CRC32;			

Query Match 91.4%; Score 32; DB 1; Length 235;  
Best Local Similarity 85.7%; Pred. No. 6.8;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

```

Db      59 KFLKFL 65
|||||
RESULT 2
HGHI_YEAST      STANDARD;      PRT;      394 AA.
AC      P48362;
DT      01-FEB-1996 (REL. 33, CREATED)
DT      01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE      HGHI PROTEIN.
GN      HGHI OR YGR187C OR G7538.
OS      SACCHAROMYCES CREVISIAE (BAKER'S YEAST).
OC      EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
CC      SACCHAROMYCETACEAE; SACCHAROMYCES.
RN      [1]
RP      SEQUENCE OF 111-394 FROM N.A.
RA      SUN Z., LIANG J., HAMPESEY M.;
RL      SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C;
RX      MEDLINE; 97279231.
RA      ARROYO J., GARCIA-GONZALEZ M., GARCIA-SAEZ M.I., SANCHEZ-PEREZ M.,
RA      NOMBELA C.;
RT      "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT      of Saccharomyces cerevisiae chromosome VII.";
RL      YEAST 13:357-363(1997).
CC      -!- SIMILARITY: TO S.POMBE SPAC26F1.12C.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U27358; G1002516; -
DR      EMBL; 272972; E243565; -
DR      EMBL; X99074; E252631; -
DR      SGD; L0003003; HGHI.
FT      DOMAIN 371 389      ASP/GLU-RICH (HIGHLY ACIDIC).
SQ      SEQUENCE 394 AA; 44951 MW; EE1B34E3 CRC32;

Query Match      85.7%; Score 30; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KFLKFL 6
Db      91 KFLKFL 96
|||||
RESULT 3
PMAL_LYCES      STANDARD;      PRT;      956 AA.
AC      P22180;
DT      01-AUG-1991 (REL. 19, CREATED)
DT      01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT      01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE      PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).
GN      LHAI
OS      LYCOPERSICON ESCULENTUM (TOMATO).
OC      EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC      EUPHYLOPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC      ASTERIDAE; SOLANACEAE; SOLANACEAE; SOLANACEAE; SOLANUM.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. CASTELMART; TISSUE=ROOT;
RA      EWING N.N., WIMMERS L.E., MEYER D.J., CHETELAT R.T., BENNETT A.B.;

"molecular cloning of tomato plasma membrane H+-ATPase.";
PLANT PHYSIOL. 94:1874-1881(1990).
-!- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A
HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE
ACTIVE TRANSPORT OF NUTRIENTS BY H+-SYMPORT. THE RESULTING
EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKALIZATION MAY MEDIATE
GROWTH RESPONSES.
-!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
-!- SUBUNIT: POSSIBLY EXISTS AS AN HOMODIMER OR AN HOMOTRIMER.
-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-!- AS MANY AS 6 TO 8 CLOSELY RELATED GENES MAY ENCODE OTHER ISOFORMS
OF PLASMA MEMBRANE ATPASE IN TOMATO, LIKE THE LHAI GENE PRODUCT
WHICH IS 96% IDENTICAL TO THE LHAI GENE PRODUCT.
-!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
(EL-E2 ATPASES).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M60166; G170464; -
DR      PIR; A45506; A45506.
DR      PROSITE; PS00154; ATPASE_EL_E2; 1.
DR      PFAM; PF00122; EL-E2_ATPase; 1.
KW      HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;
KW      ATP-BINDING; MULTIGENE FAMILY.
FT      DOMAIN 1 64      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 65 85      1 (POTENTIAL).
FT      DOMAIN 86 96      CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 97 116      2 (POTENTIAL).
FT      DOMAIN 117 245      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 246 267      3 (POTENTIAL).
FT      DOMAIN 268 279      CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 280 302      4 (POTENTIAL).
FT      DOMAIN 303 677      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 678 696      5 (POTENTIAL).
FT      DOMAIN 697 712      CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 713 736      6 (POTENTIAL).
FT      DOMAIN 737 794      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 795 815      7 (POTENTIAL).
FT      DOMAIN 816 830      CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 831 847      8 (POTENTIAL).
FT      DOMAIN 848 956      EXTRACELLULAR (POTENTIAL).
FT      MOD_RES 333 333      PHOSPHORYLATION (BY SIMILARITY).
FT      BINDING 427 427      ATP (BY SIMILARITY).
SQ      SEQUENCE 956 AA; 105103 MW; E798A2DA CRC32;

Query Match      85.7%; Score 30; DB 1; Length 956;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KFLKFL 6
Db      61 KFLKFL 66
|||||
RESULT 4
RR6_CYACA
ID      RR6_CYACA      STANDARD;      PRT;      103 AA.
AC      O19917;
DT      15-JUL-1998 (REL. 36, CREATED)
DT      15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT      15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE      CHLOROPLAST 30S RIBOSOMAL PROTEIN S6.
GN      RPS6.
OS      CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OC      EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDACEAE;
OG      CHLOROPLAST.

```

OC CYANIDIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RK-1;  
RA GLOCKNER G., ROSENTHAL A., VALENTIN K.;  
RT "Organisation of 46 kb of the Cyanidium caldarium RK1 plastid genome";  
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: BINDS TOGETHER WITH S18 TO 16S RIBOSOMAL RNA (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF022186; G2465743; -  
DR PROSITE; PS01048; RIBOSOMAL\_S6; FALSE\_NEG.  
DR PFAM; PF01250; Ribosomal\_S6; 1.  
KW RIBOSOMAL PROTEIN; RNA-BINDING; CHLOROPLAST.  
SQ SEQUENCE 103 AA; 12340 MW; CC91D65A CRC32;  
  
Query Match 85.7%; Score 30; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 7.6; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KFLKFL 6  
Db 28 KFLKFL 33  
  
RESULT 5  
Y940\_METJA ID Y940\_METJA STANDARD; PRT; 318 AA.  
AC Q58350;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN MJ0940.  
GN MJ0940.  
OS METHANOCOCCUS JANNASCHII.  
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
CC METHANOCOCCUS.  
CC  
CC SEQUENCE FROM N.A.  
RN STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE; 96337999.  
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLOBE A.,  
RA SCOTT J.L., GEORGE N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii";  
RL SCIENCE 273:1058-1073(1996).  
CC -1- SIMILARITY: CONTAINS ? TPR DOMAINS.  
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0941.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL; U67537; G1591608; -  
DR TIGR; MJ0940; -  
DR PFAM; PF00515; TPR; 2  
KW HYPOTHETICAL PROTEIN; REPEAT; TPR DOMAIN.  
SQ SEQUENCE 318 AA; 36861 MW; 7779AE60 CRC32;  
  
Query Match 85.7%; Score 30; DB 1; Length 318;  
Best Local Similarity 85.7%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KFLKFL 7  
Db 51 KFAKFL 57  
  
RESULT 6  
CYSF\_PLAFA ID CYSF\_PLAFA STANDARD; PRT; 569 AA.  
AC P25805;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE THROPHOZOITE CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-) (TCP).  
OS PLASMODIUM FALCIPARUM.  
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92228005.  
RA ROSENTHAL P.J., NELSON R.G.;  
RT "Isolation and characterization of a cysteine proteinase gene of Plasmodium falciparum";  
RL MOL. BIOCHEM. PARASITOL. 51:143-152(1992).  
CC -1- FUNCTION: PROBABLY DEGRADES ERTHROCYTE HEMOGLOBIN.  
CC -1- DEVELOPMENTAL STAGE: THROPHOZOITE.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE PAPAIN FAMILY OF THIOL PROTEASES.  
CC -1- SIMILARITY: STRONGEST SIMILARITY TO CATHEPSIN L.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M81341; G160248; -  
DR PIR; A45624; A45624.  
DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
DR PFAM; PF00112; Cys-protease; 2.  
DR HSSP; P07858; 1CSB.  
KW HYDROLASE; THIOL PROTEASE; ZYMOGEN; GLYCOPROTEIN; SIGNAL.  
FT SIGNAL 1 ?  
FT PROPEP ? 332  
FT CHAIN 333 569  
FT DOMAIN 64 70  
FT ACT\_SITE 357 357  
FT ACT\_SITE 488 488  
FT ACT\_SITE 533 533  
FT ACT\_SITE 534 395  
FT DISULFID 58 58  
FT CARBOHYD 98 98  
FT CARBOHYD 121 121  
FT CARBOHYD 127 127  
FT CARBOHYD 479 479  
FT CARBOHYD 487 487  
SQ SEQUENCE 569 AA; 66880 MW; 9B8F0096 CRC32;

Query Match 82.9%; Score 29; DB 1; Length 569;  
 Best Local Similarity 71.4%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 224 KFFRFMK 230

RESULT 7  
 FLAE\_METJA STANDARD; PRT; 140 AA.  
 ID Q58306;  
 AC Q58306;  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE PUTATIVE FLAGELLA-RELATED PROTEIN E.  
 GN FLAE OR M0896.  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 CC METHANOCOCCUS.  
 RN [1]

SEQUENCE FROM N.A.  
 RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,  
 RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL SCIENCE 273:1058-1073(1996).  
 CC -!- SIMILARITY: STRONG, TO M.VOLTAE FLAE, ALSO TO FLAD.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U67533; G1591572; -;  
 DR TIGR; MJ0896; -;  
 KW FLAGELLA.  
 SQ SEQUENCE 140 AA; 15891 MW; 61446FEB CRC32;

Query Match 82.9%; Score 29; DB 1; Length 140;  
 Best Local Similarity 85.7%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 73 KLLKFLK 79

RESULT 8  
 HXTC\_YEAST STANDARD; PRT; 564 AA.  
 ID HXTC\_YEAST  
 AC P39924;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DE HEXOSE TRANSPORTER HXT13.  
 GN HXT13 OR HXT8 OR YEL069C.  
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
 CC SACCCHAROMYCETACEAE; SACCCHAROMYCES.

SEQUENCE FROM N.A.  
 RP STRAIN-S288C / AB972;  
 RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
 RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,  
 RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKER-SMITH S.,  
 RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,  
 RA MOSDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,  
 RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOENEN T., SMITH V.,  
 RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;  
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U18795; G603249; -;  
 DR SGD; L0002640; HXT13.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 DR PFAM; PF000083; sugar.tr; 1.  
 KW REPEAT; TRANSMEMBRANE; SUGAR TRANSPORT; TRANSPORT; GLYCOPROTEIN.  
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 52 73 POTENTIAL.  
 FT DOMAIN 74 107 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 108 127 POTENTIAL.  
 FT DOMAIN 128 136 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 137 156 POTENTIAL.  
 FT DOMAIN 157 163 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 164 185 POTENTIAL.  
 FT DOMAIN 186 197 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 198 217 POTENTIAL.  
 FT DOMAIN 218 229 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 230 250 POTENTIAL.  
 FT DOMAIN 251 318 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 319 339 POTENTIAL.  
 FT DOMAIN 340 352 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 353 374 POTENTIAL.  
 FT DOMAIN 375 381 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 382 401 POTENTIAL.  
 FT DOMAIN 402 421 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 422 445 POTENTIAL.  
 FT DOMAIN 446 457 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 458 479 POTENTIAL.  
 FT DOMAIN 480 484 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 485 506 POTENTIAL.  
 FT DOMAIN 507 564 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 564 AA; 62734 MW; 073CAE7E CRC32;

Query Match 82.9%; Score 29; DB 1; Length 564;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7  
 DB 555 KFLKFSK 561

RESULT 9  
 HXTC\_YEAST STANDARD; PRT; 564 AA.  
 ID HXTC\_YEAST  
 AC P53631;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DE HEXOSE TRANSPORTER HXT17.  
GN HXT17 OR YNR072W OR N3615.  
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCAROMYCETALES;  
CC SACCAROMYCETACEAE; SACCAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDRE B., IRAQI ROUSSAINI I., URRESTARAZU L.A., VISSERS S.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; 271687; E239602; -  
CC SGD; L0003265; HXT17.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
DR PFAM; PF00083; sugar\_tr; 1.  
DR REPEAT; TRANSMEMBRANE; SUGAR TRANSPORT; TRANSPORT; GLYCOPROTEIN.  
FT TRANSMEM 50 70  
FT TRANSMEM 110 130 POTENTIAL.  
FT TRANSMEM 137 157 POTENTIAL.  
FT TRANSMEM 168 188 POTENTIAL.  
FT TRANSMEM 196 216 POTENTIAL.  
FT TRANSMEM 230 250 POTENTIAL.  
FT TRANSMEM 318 338 POTENTIAL.  
FT TRANSMEM 353 373 POTENTIAL.  
FT TRANSMEM 382 402 POTENTIAL.  
FT TRANSMEM 423 443 POTENTIAL.  
FT TRANSMEM 463 483 POTENTIAL.  
FT TRANSMEM 485 505 POTENTIAL.  
SQ SEQUENCE 564 AA; 62828 MW; 23C1F139 CRC32;  
-----  
Query Match 82.9%; Score 29; DB 1; Length 564;  
Best Local Similarity 85.7%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7  
DB 555 KFLKFSK 561  
-----  
RESULT 10  
RRPB\_BEV  
ID RRPB\_BEV STANDARD; PRT; 2291 AA.  
AC P18458;  
DT 01-NOV-1990 (REL. 16, CREATED)  
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (ORF1b).  
GN POL.  
OS BERNE VIRUS (BEV).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; NIDOVIRALES;  
OC CORONAVIRIDAE; TOROVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE P138/72;  
RX MEDLINE; 90356389.  
RA SNIJDER E.J., DEN BOON J.A., BREDEBEEK P.J., HORZINEK M.C.,  
RA RIJNBRAND R., SPAAN W.J.M.;  
RT "The carboxyl-terminal part of the putative Berne virus polymerase is  
RT expressed by ribosomal frameshifting and contains sequence motifs  
RT which indicate that toro- and coronaviruses are evolutionarily  
RT related."  
-----

RL NUCLEIC ACIDS RES. 18:4535-4542(1990).  
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE  
CC + RNA(N).  
CC -1- THIS PROTEIN IS EXPRESSED BY AN EFFICIENT RIBOSOMAL FRAMESHIFTING  
CC MECHANISM. RIBOSOMAL FRAMESHIFTING IS AN ELEGANT MECHANISM FOR  
CC REGULATING THE SYNTHESIS OF SEVERAL PROTEINS IN A WELL BALANCED  
CC MANNER.  
CC -1- SIMILARITY: TO CORONAVIRUSES POLYMERASE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X52374; E23105; -  
CC PIR; S11238; S11238.  
KW TRANSFERASE; RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
FT DOMAIN 515 744 POLYMERASE.  
FT DOMAIN 1099 1374 HELICASE.  
FT NP\_BIND 1101 1108 ATP (BY SIMILARITY).  
FT SIMILAR 845 928 TO ITS HOMOLOG FROM MHV (AA 938-1022)  
FT SIMILAR 1921 2209 AND IBV (AA 929-1013).  
FT SIMILAR TO ITS HOMOLOG FROM MHV (AA 2322-2641)  
FT AND IBV (AA 2274-2598).  
SQ SEQUENCE 2291 AA; 260843 MW; C0283C2D CRC32;  
-----  
Query Match 82.9%; Score 29; DB 1; Length 2291;  
Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFLKFLK 7  
DB 524 KFSKFLK 530  
-----  
RESULT 11  
RT07\_ACACA  
ID RT07\_ACACA STANDARD; PRT; 337 AA.  
AC P46756;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.  
GN RPS7.  
OS ACANTHAMOEBA CASTELLANII (AMOEB).  
OC MITOCHONDRION.  
OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 30010 / NEFF;  
RX MEDLINE; 93157849.  
RA LONERGAN K.M., GRAY M.W.;  
RT "Editing of transfer RNAs in Acanthamoeba castellanii mitochondria."  
RT SCIENCE 259:812-816(1993).  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U12386; G562050; -  
CC PROSITE; PS00052; RIBOSOMAL\_S7; FALSE\_NEG.  
CC PFAM; PF00177; S7; 1.  
KW RIBOSOMAL PROTEIN; MITOCHONDRION.  
-----

SQ SEQUENCE 337 AA; 41612 MW; 9817FLE4 CRC32;

Query Match 82.9%; Score 29; DB 1; Length 337;

Best Local Similarity 71.4%; Pred. No. 39;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

|||:||||

Db 328 RFLRFLK 334

RESULT 12

Y011\_MYCGE

AC P47257; Q49362; STANDARD; PRT; 287 AA.

DT 01-FEB-1996 (REL. 33, CREATED)

DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN MG011.

GN MG011.

OS MYCOPLASMA GENITALIUM.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;

OC MYCOPLASMATACEAE; MYCOPLASMA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 33530 / G-37;

RX MEDLINE; 96026346.

RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,

RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,

RA FRITZCHMANN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,

RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,

RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,

RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;

RT "The minimal gene complement of Mycoplasma genitalium.";

RL SCIENCE 270:397-403(1995).

RN [2]

RP SEQUENCE OF 158-255 FROM N.A.

RC STRAIN-ATCC 33530 / G-37;

RX MEDLINE; 94075230.

RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

sequencing.";

RL J. BACTERIOL. 175:7918-7930(1993).

CC -!- SIMILARITY: BELONGS TO THE ATP-GRASP SUPERFAMILY; RIMK FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U39679; G1045679; -.

DR EMBL; U02257; G406920; -.

DR TIGR; MG011; -.

KW HYPOTHETICAL PROTEIN.

FT CONFLICT 158 158 T -> P (IN REF. 2).

FT CONFLICT 255 255 A -> R (IN REF. 2).

SQ SEQUENCE 287 AA; 33433 MW; 9E2DA74F CRC32;

Query Match 82.9%; Score 29; DB 1; Length 287;

Best Local Similarity 85.7%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

|||:||||

Db 276 KFKKFLK 282

RESULT 13

Y343\_MYCGE

ID Y343\_MYCGE STANDARD; PRT; 346 AA.

AC P47585;

DT 01-FEB-1996 (REL. 33, CREATED)

DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN MG343.

GN MG343.

OS MYCOPLASMA GENITALIUM.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;

OC MYCOPLASMATACEAE; MYCOPLASMA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 33530 / G-37;

RX MEDLINE; 96026346.

RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,

RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,

RA FRITZCHMANN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,

RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,

RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,

RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;

RT "The minimal gene complement of Mycoplasma genitalium.";

RL SCIENCE 270:397-403(1995).

RN [2]

RP SEQUENCE OF 37-150 FROM N.A.

RC STRAIN-ATCC 33530 / G-37;

RX MEDLINE; 94075230.

RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

sequencing.";

RL J. BACTERIOL. 175:7918-7930(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U39718; G1046049; -.

DR EMBL; U01811; G407662; -.

DR TIGR; MG343; -.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 346 AA; 40834 MW; 1450B76F CRC32;

Query Match 82.9%; Score 29; DB 1; Length 346;

Best Local Similarity 85.7%; Pred. No. 41;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFLK 7

|||||

Db 112 KFLKILK 118

RESULT 14

BEM2\_YEAST

ID BEM2\_YEAST STANDARD; PRT; 2167 AA.

AC P39960;

DT 01-FEB-1995 (REL. 31, CREATED)

DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE GTPASE ACTIVATING PROTEIN BEM2/IPL2.

GN BEM2 OR IPL2 OR SUP9 OR YER155C.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

OC SACCHAROMYCETACEAE; SACCHAROMYCES.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-S288C.

RX MEDLINE; 95050996.

RA KIM Y., FRANCISCO L., CHEN G., MARCOTTE E., CHAN C.S.;



RT "Control of cellular morphogenesis by the Ipl2/Bem2 GTPase-activating  
 RL protein: possible role of protein phosphorylation."  
 RN J. CELL BIOL. 127:1381-1394(1994).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 95050997.  
 RA PETERSON J., ZHENG Y., BENDER L., MYERS A., CERIONE R., BENDER A.:  
 "Interactions between the bud emergence proteins Bem1p and Bem2p and  
 RL Rho-type GTPases in yeast."  
 RN J. CELL BIOL. 127:1395-1406(1994).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
 RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,  
 RA CHUNG R., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKE-SMITH S.,  
 RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,  
 RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,  
 RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGREN T., SMITH V.,  
 RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.:  
 RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE: 94123964.  
 RA CHAN C.S., BOTSTEIN D.:  
 "Isolation and characterization of chromosome-gain and increase-in-  
 RL ploidy mutants in yeast."  
 RT GENETICS 135:677-691(1993).  
 CC INVOLVED IN THE CONTROL OF CELLULAR MORPHOGENESIS. REQUIRED FOR  
 CC PROPER BUD SITE SELECTION AND BUD EMERGENCE.  
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 GAP DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; 235159; G51137; -  
 DR EMBL; L33832; G49695; -  
 DR EMBL; U18917; G60395; -  
 DR SGD; L0000168; BEM2.  
 DR PROSITE; PS00003; PH DOMAIN; 1.  
 DR PFAM; PF00617; RasGEF; 1.  
 DR PFAM; PF00618; RasGEF; 1.  
 DR PFAM; PF00620; RhoGAP; 1.  
 KW GTPASE ACTIVATION.  
 FT DOMAIN 16 25 POLY-SER.  
 FT DOMAIN 35 43 POLY-SER.  
 FT DOMAIN 58 63 POLY-HIS.  
 FT DOMAIN 198 208 POLY-ASN.  
 FT DOMAIN 253 260 POLY-SER.  
 FT DOMAIN 1161 1165 POLY-THR.  
 FT DOMAIN 1914 1948 PH.  
 FT DOMAIN 1981 2134 GAP DOMAIN.  
 SQ SEQUENCE 2167 AA; 245428 MW; DA2FF455 CRC32;  
 Query Match 80.0%; Score 28; DB 1; Length 2167;  
 Best Local Similarity 83.3%; Pred. No. 4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 FLKFLK 7  
 Db 463 FIKFLK 468  
 RESULT 15  
 DIP2\_YEAST

ID DIP2\_YEAST STANDARD; PRT; 943 AA.  
 AC Q12220; Q05386;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE DOM34 INTERACTING PROTEIN 2.  
 GN DIP2 OR YLR129W OR L9233.1 OR L3116.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
 RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
 RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
 RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
 RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
 RA TAILCH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
 RA WILSON R., WATERSTON R.:  
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RA DELIUS H.:  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [3]  
 RP SEQUENCE OF 1-458 FROM N.A.  
 RC STRAIN-FY23 / RD005;  
 RA VERHASSELT P., VOET M., VOLCKAERT G.:  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: TO S.POMBE SPC3D5.12.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U53877; G1256869; -  
 DR EMBL; U53881; G2580465; -  
 DR EMBL; X91258; G995691; -  
 DR EMBL; Z73301; E245568; -  
 DR EMBL; Z73302; E322129; -  
 DR EMBL; X89514; E198829; -  
 DR SGD; L0003088; DIP2.  
 DR PROSITE; PS00678; WD\_REPEATS; 4.  
 DR PFAM; PF00400; G-beta; 8.  
 KW REPEAT; WD REPEAT.  
 FT REPEAT 77 107 WD1.  
 FT REPEAT 119 149 WD2.  
 FT REPEAT 161 190 WD3.  
 FT REPEAT 202 230 WD4.  
 FT REPEAT 389 418 WD5.  
 FT REPEAT 428 458 WD6.  
 FT REPEAT 471 501 WD7.  
 FT REPEAT 571 601 WD8.  
 FT REPEAT 613 643 WD9.  
 FT REPEAT 655 685 WD10.  
 SQ SEQUENCE 943 AA; 106342 MW; 7B348EA8 CRC32;  
 Query Match 80.0%; Score 28; DB 1; Length 943;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KFLKFL 6  
 Db 831 KFLKFI 836

Search completed: September 7, 1999, 23:59:18  
Job time: 530 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:48 ; Search time 160.2 Seconds  
(without alignments)  
2.689 Million cell updates/sec

Title: US-09-124-280A-38

Perfect score: 35

Sequence: 1 KFLKFLK 7

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	32	91.4	140	2	032317	032317 bacillus th
2	32	91.4	642	5	018203	018203 caenorhabdi
3	31	88.6	395	1	058134	058134 pyrococcus
4	31	88.6	350	5	019556	019556 caenorhabdi
5	30	85.7	260	1	027551	027551 methanobact
6	30	85.7	288	2	051841	051841 porphyromon
7	30	85.7	141	2	067910	067910 aquifex aeo
8	30	85.7	53	2	068585	068585 streptococc
9	30	85.7	493	2	067365	067365 aquifex aeo
10	30	85.7	193	2	066417	066417 aquifex aeo
11	30	85.7	141	2	005963	005963 rickettsia
12	30	85.7	309	3	099188	099188 saccharomyc
13	30	85.7	1462	3	042874	042874 schizosacch
14	30	85.7	640	3	093831	093831 candida alb
15	30	85.7	342	4	098633	098633 homo sapien
16	30	85.7	127	5	016223	016223 caenorhabdi
17	30	85.7	1365	5	077308	077308 plasmodium
18	30	85.7	1817	5	096253	096253 plasmodium
19	30	85.7	951	10	043106	043106 phaseolus v
20	30	85.7	953	10	093265	093265 mesembryant
21	30	85.7	1531	10	004565	004565 arabidopsis
22	30	85.7	956	10	042932	042932 nicotiana p
23	30	85.7	963	10	043131	043131 vicia faba
24	30	85.7	956	10	043182	043182 solanum tub
25	30	85.7	949	10	043243	043243 zea mays (m
26	30	85.7	956	10	043001	043001 oryza sativ
27	30	85.7	957	10	043002	043002 oryza sativ
28	29	82.9	312	1	058122	058122 pyrococcus
29	29	82.9	1279	2	026046	026046 helicobacte

## ALIGNMENTS

RESULT 1  
032317 PRELIMINARY; PRT; 140 AA.  
AC 032317;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 15 9 KD PROTEIN.  
OS Bacillus thuringiensis.  
OG Plasmid pgi3.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
RN [1]  
RX MEDLINE; 97405895.  
RA HOFACK L., SEURINCK J., MAHILLON J.;  
RT "Nucleotide sequence and characterization of the cryptic Bacillus  
RT thuringiensis plasmid pgi3 reveal a new family of rolling circle  
RT replicons."  
RL J. Bacteriol. 179:5000-5008(1997).  
DR EMBL; Y11173; CAA72050.1; -.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 140 AA; 15901 MW; 48971A3F CRC32;

Query Match 91.4%; Score 32; DB 2; Length 140;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KFLKFLK 7  
DB 4 KFKVFLK 10  
||:||||  
|:|:|

## RESULT 2

018203 PRELIMINARY; PRT; 642 AA.  
AC 018203;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE Y48E1B.13 PROTEIN.  
GN Y48E1B.13  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MCMURRAY A.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

```

RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P., III of C.
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z93393; CAB07698.1; -.
DR PFAM; PF00655; ICE_P10; 1.
DR PFAM; PF00656; ICE_P20; 1.
SQ SEQUENCE 642 AA; 73590 MW; B335IDE0 CRC32;

Query Match 91.4%; Score 32; DB 5; Length 642;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7
Db 157 RFLKFLK 163

RESULT 3
O58134 PRELIMINARY; PRT; 395 AA.
AC O58134;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE 395AA LONG HYPOTHETICAL PROTEIN.
GN P0397.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE; 98344137.
RA KAWARABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
RA KIKUCHI H.;
RT "Complete sequence and gene organization of the genome of a
RT hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998)
DR EMBL; AP000002; BAA29472.1; -.
SQ SEQUENCE 395 AA; 46026 MW; D2DC6575 CRC32;

Query Match 88.6%; Score 31; DB 1; Length 395;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7
Db 237 KFLKFLK 243

RESULT 4
Q19556 PRELIMINARY; PRT; 350 AA.
AC Q19556;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)

```

```

DE F18E2.4 PROTEIN.
GN F18E2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA LIGHTNING J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P., III of C.
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z75537; CAA9837.1; -.
SQ SEQUENCE 350 AA; 39851 MW; 9AA64F5B CRC32;

Query Match 88.6%; Score 31; DB 5; Length 350;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7
Db 320 EFLKFLK 326

RESULT 5
O27551 PRELIMINARY; PRT; 260 AA.
AC O27551;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
DE 2-HYDROXYHEPTA-2,4-DIENE-1,7-DIOLATE ISOMERASE.
GN MTH1507.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000911; AAB85982.1; -.
KW Isomerase.
SQ SEQUENCE 260 AA; 28873 MW; 536811AB CRC32;

Query Match 85.7%; Score 30; DB 1; Length 260;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy 1 KFLKFLK 7
    |||:|:|
Db 5 KFLRFMK 11

RESULT 6
Q51841 PRELIMINARY; PRT; 288 AA.
ID Q51841
AC Q51841;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE GPDXJ, PUT, AND YHEG-PG GENES.
GN Put.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Cytophagales; Bacteroidaceae; Porphyromonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W83;
RA WALLACE A.M., ROBERTS I.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X97228; CAA65865.1; -.
SQ SEQUENCE 288 AA; 31805 MW; 0460290B CRC32;

Query Match 85.7%; Score 30; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFLKFLK 7
    ||| |||
Db 38 KFLSFLK 44

RESULT 7
O67910 PRELIMINARY; PRT; 141 AA.
ID O67910
AC O67910;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 16.6 KD PROTEIN.
GN AQ_2157.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE000776; AAC07885.1; -.
KW Hypothetical protein.
SQ SEQUENCE 141 AA; 16618 MW; E880929A CRC32;

Query Match 85.7%; Score 30; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFLKFLK 6
    |||||
Db 5 KFLRFMK 11

RESULT 8
O68585 PRELIMINARY; PRT; 53 AA.
ID O68585
AC O68585;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 6.2 KD PROTEIN.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JH1005;
RX MEDLINE; 98261521.
RA HILLMAN J.D., NOVAK J., SAGURA E., GUTIERREZ J.A., BROOKS T.A.,
RA CROWLEY P.J., HESS M., AZIZI A., LEUNG K.-P., CVITKOVITCH D.,
RA BLEIWEIS A.S.;
RT "Genetic and biochemical analysis of mutacin 1140, a lantibiotic from
RT Streptococcus mutans."
RL Infect. Immun. 66:2743-2749(1998).
DR EMBL; AF051560; AAC18826.1; -.
KW Hypothetical protein.
SQ SEQUENCE 53 AA; 6228 MW; C2D60052 CRC32;

Query Match 85.7%; Score 30; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLKFLK 7
    |||||
Db 6 FLKFLK 11

RESULT 9
O67365 PRELIMINARY; PRT; 493 AA.
ID O67365
AC O67365;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 56.5 KD PROTEIN.
GN AQ_1350.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE000736; AAC07337.1; -.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 56478 MW; E1603544 CRC32;

Query Match 85.7%; Score 30; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
|||||

Db 188 KFLKFL 193

RESULT 10

O66417 PRELIMINARY; PRT; 193 AA.

AC O66417;

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE HYPOTHETICAL 22.6 KD PROTEIN.

GN AA26

OS Aquifex aeolicus.

OG Plasmid ecel.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 98196666.

RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,

RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,

RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus";

RL Nature 392:353-358(1998).

DR EMBL; AF000667; AAC07989.1; -.

KW Hypothetical protein; Plasmid.

SQ SEQUENCE 193 AA; 22552 MW; 935BAF96 CRC32;

Query Match

Best Local Similarity 85.7%; Score 30; DB 2; Length 193;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7

|||||

Db 57 FLKFLK 62

RESULT 11

O05963

ID O05963 PRELIMINARY; PRT; 141 AA.

AC O05963;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE HYPOTHETICAL 16.5 KD PROTEIN.

GN RP436.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MADRID E;

RX MEDLINE; 97419517.

RA ANDERSSON J.O., ANDERSSON S.G.E.;

RT "Genomic rearrangements during evolution of the obligate

intracellular parasite Rickettsia prowazekii as inferred from an

analysis of 52015 bp nucleotide sequence.";

RL Microbiology 143:2783-2795(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-MADRID E;

RX MEDLINE; 99039499.

RA ANDERSSON S.G.E., ZOMORODIPOUR A., ANDERSSON J.O.,

RA ERIKSSON A.S., WINKLER H.H., KURLAND C.G.;

RT "The genome sequence of Rickettsia prowazekii and the origin of

mitochondria.";

RL Nature 396:133-140(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-MADRID E;

RA ANDERSSON S.G.E.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y11777; CAA72442.1; -.

DR EMBL; AJ235271; CAA14893.1; -.

KW Hypothetical protein.

SQ SEQUENCE 141 AA; 16493 MW; 8AF2457A CRC32;

Query Match

Best Local Similarity 85.7%; Score 30; DB 2; Length 141;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6

|||||

Db 118 KFLKFL 123

RESULT 12

Q99188

ID Q99188 PRELIMINARY; PRT; 309 AA.

AC Q99188;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)

DE ORF YOR107W FROM CHROMOSOME XV.

GN YOR107W AND YOR3224W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;

OC Saccharomycetaceae; Saccharomycetes.

RN [1]

RP SEQUENCE FROM N.A.

RA BENES V., ANDRADE M.A., RECHMANN S., TEODORU C., BANREVI A.,

RA SANDER C., VALENCIA A., ANSORGE W., VOSS H.;

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA VOSS H., BENES V., RECHMANN S., TEODORU C., SCHWAGER C., PACES V.,

RA ANSORGE W.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA MIPS;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; X94335; CAA64027.1; -.

DR EMBL; Z75015; CAA99305.1; -.

SQ SEQUENCE 309 AA; 35100 MW; 04E1C8F9 CRC32;

Query Match

Best Local Similarity 85.7%; Score 30; DB 3; Length 309;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFLKFL 7

|||||

Db 50 KFYKFLK 56

RESULT 13

O42874

ID O42874 PRELIMINARY; PRT; 1462 AA.

AC O42874;

DT 01-JAN-1999 (TREMBLrel. 09, Created)

DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)

DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)

DE HYPOTHETICAL 164.1 KD PROTEIN C3G9.12 IN CHROMOSOME 1.

GN SPAC3G9.12.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

RN SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA OLIVER K., HARRIS D., WOOD V., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO YEAST STU1.  
 DR EMBL; AL021046; CAA15921.1; -;  
 KW Hypothetical protein.  
 FT DOMAIN 226 229 POLY-SER.  
 FT DOMAIN 855 861 POLY-SER.  
 SQ SEQUENCE 1462 AA; 164093 MW; FF20F54E CRC32;

Query Match 85.7%; Score 30; DB 3; Length 1462;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLKFLK 7  
 DB 9 FLKFLK 14  
 |||||

RESULT 14  
 O93831 PRELIMINARY; PRT; 640 AA.  
 AC O93831;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE RAB GERANYLGERANYL TRANSFERASE ESCORT PROTEIN (REP).  
 GN MRS6.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Candidaceae; Candida.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-IFO1060;  
 RA ISHII N., AOKI Y., ARISAWA M.;  
 RT "Molecular cloning of Rab geranylgeranyl transferase escort protein  
 (REP) homologue from Candida albicans";  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB021317; BAA36167.1; -;  
 KW Transferase.  
 SQ SEQUENCE 640 AA; 72291 MW; C54E8D44 CRC32;

Query Match 85.7%; Score 30; DB 3; Length 640;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
 DB 158 KFLKFL 163  
 |||||

RESULT 15  
 Q99633 PRELIMINARY; PRT; 342 AA.  
 AC Q99633;  
 DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE HRP18.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE; 97152474.  
 RA HOROWITZ D.S., KRAINER A.R.;  
 RT "A human protein required for the second step of pre-mRNA splicing is  
 RT functionally related to a yeast splicing factor.";  
 RL Genes Dev. 11:139-151(1997).

DR EMBL; U51990; AAB41490.1; -;  
 SQ SEQUENCE 342 AA; 39859 MW; BFD392F8 CRC32;  
 Query Match 85.7%; Score 30; DB 4; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFLKFL 6  
 DB 190 KFLKFL 195  
 |||||

Search completed: September 7, 1999, 22:47:49  
 Job time: 7969 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:08 ; Search time 135.78 Seconds  
(without alignments)  
1.570 Million cell updates/sec

Title: US-09-124-280A-39

Perfect score: 48

Sequence: 1 RYRVRYRV 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	9	1 R33534	Peptide for treati
2	48	100.0	9	1 R39296	Endotoxin lipid A
3	48	100.0	9	1 R71780	Peptide neutralisi
4	48	100.0	9	1 W21627	Antibiotic potenti
5	48	100.0	9	1 W21597	Antibiotic potenti
6	36	75.0	15	1 W22306	Immunoglobulin-bi
7	35	72.9	463	1 W89250	Mouse PTP05 isofor
8	35	72.9	405	1 W89251	Mouse PTP05 isofor
9	35	72.9	122	1 W89252	Rat PTP10. New nuc
10	35	72.9	426	1 W89249	Mouse PTP05. New n
11	34	70.8	394	1 W88313	O antigen polymera
12	32	66.7	6	1 R37386	Peptide for treati
13	31	64.6	1422	1 R82070	Hepatitis GB virus
14	30	62.5	553	1 R13944	Partial HVT ribonu
15	30	62.5	6	1 R37387	Peptide for treati
16	30	62.5	1013	1 R62535	P. vulgaris chondr
17	30	62.5	990	1 R84660	Chondroitinase II.
18	30	62.5	990	1 R77460	Chondroitinase-II.
19	30	62.5	568	1 W03325	Newcastle disease
20	30	62.5	568	1 W03138	Newcastle disease
21	30	62.5	1013	1 W09398	Proteus vulgaris C
22	30	62.5	568	1 W03551	Newcastle disease
23	30	62.5	990	1 W39746	P. vulgaris chondr
24	30	62.5	990	1 W37786	Amino acid sequenc
25	30	62.5	144	1 W55084	Streptococcus pneu
26	30	62.5	568	1 W62989	HN protein of Newc
27	30	62.5	571	1 W44940	Newcastle disease
28	30	62.5	568	1 W49683	Newcastle disease
29	30	62.5	30	1 W74846	Human secreted pro
30	30	62.5	1118	1 W82395	Human UBP protein
31	30	62.5	990	1 W90075	P. vulgaris chondr
32	30	62.5	194	1 Y11220	S. pneumoniae prot
33	29	60.4	448	1 R14478	51.4 kD toxin. Cau
34	29	60.4	554	1 R26271	Alpha-acetolactate
35	29	60.4	241	1 R39515	Caffeoyl-CoA-3-O-m
36	29	60.4	448	1 R41016	Insecticidal prote
37	29	60.4	448	1 R49873	Bacillus sphaericu
38	29	60.4	377	1 R56529	Protein kinase (CK
39	29	60.4	347	1 R56530	Protein kinase (CK
40	29	60.4	377	1 R76625	Human HRR25-like c
41	29	60.4	347	1 R76626	Human HRR25-like c
42	29	60.4	98	1 W55298	H. pylori ORF 06cp
43	29	60.4	134	1 W55658	H. pylori ORF 11ae

ALIGNMENTS

RESULT 1

```
R33534
ID R33534 standard; peptide; 9 AA.
AC R33534:
DT 07-JUL-1993 (first entry)
DE Peptide for treating septic shock.
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;
KW cytokine release control; treatment; pertussis; bacterial meningitis;
KW HIV related infections; polymyxin B.
OS Synthetic.
PN ZA9200943-A.
PD 25-NOV-1992.
PF 10-FEB-1992; 000943.
PR 11-FEB-1991; US-658744.
PA (PORR/) PORRO M.
PI PORRO M;
DR WPI; 93-094304/11.
PT New peptide for treatment or prevention of toxic shock - comprises
PT specified sequences of aminoacid(s) and analogs
PT comprising sequences retro-orientated
PS Claim 12: Page 32; 39pp; English.
CC This peptide is a specific example of a generic peptide of
CC formula R-(Lys/Arg/His - Phe/Tyr/Irp - Leu/Ile/Val)n-R, where
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid
CC residue. The peptide is useful for treating or preventing septic
CC shock, mixing with polymyxin B to reduce its toxicity; removing
CC endotoxins from blood, sera or other fluids (in vivo or in
CC vitro); controlling release of cytokines induced by endotoxins;
CC as diagnostic reagents to detect and quantify toxins in blood
CC or sera; preparing non-toxic antigenic complexes of lipid A or
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial
CC meningitis and HIV-related infections. The usual dose is 10-100
CC ug/kg/day, given parenterally. It binds to the same sites as
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It
CC has no antibiotic activity; does not lyse erythrocytes; has no
CC toxicity in mice when injected at 50mg/kg and is relatively unstable
CC against proteases.
SQ Sequence 9 AA;
```

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYRVRYRV 9  
Db 1 RYRVRYRV 9

RESULT 2

```
R39296
ID R39296 standard; peptide; 9 AA.
AC R39296:
DT 22-DEC-1993 (first entry)
DE Endotoxin lipid A neutralising peptide.
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;
KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;
KW removal; sera; vaccines; drug solns.; solutions; cytokine release;
KW control; in vivo; in vitro; detoxification; detection;
OS Synthetic.
PN WO9314115-A.
PD 22-JUL-1993.
PF 14-MAY-1992; E01060.
PR 16-JAN-1992; US-819893.
PA (PORR/) PORRO M.
```

PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 12: Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solns. etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
 DB 1 RYRVRYRV 9  
 |||||

RESULT 3  
 R71780  
 ID R71780 standard; peptide; 9 AA.  
 AC R71780; 1995 (first entry)  
 DT Peptide neutralising toxicity of Lipid A.  
 DE Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 KW Synthetic.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids  
 PS Claim 11; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula:  
 CC (A)<sub>n</sub>, where A is the cationic amino acid Lys or Arg and n is 7-10;  
 CC (AB)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
 CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)<sub>p</sub>, where A is  
 CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
 CC greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
 DB 1 RYRVRYRV 9  
 |||||

W21627  
 ID W21627 standard; peptide; 9 AA.  
 AC W21627;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #39.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 43; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
 DB 1 RYRVRYRV 9  
 |||||

RESULT 5  
 W21597  
 ID W21597 standard; peptide; 9 AA.  
 AC W21597;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #9.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 13; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 48; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
 DB 1 RYRVRYRV 9  
 |||||

Qy 1 RYRVRYV 9  
 |||||  
 Db 1 RYRVRYV 9

RESULT 6  
 W12306  
 ID W12306 standard; peptide; 15 AA.  
 AC W12306;  
 DT 29-AUG-1997 (first entry)  
 DE Immunoglobulin-binding peptide used as affinity and capture ligand.  
 KW Immunoglobulin binding; affinity chromatography; immunoassay;  
 KW capture ligand; constant region; non-covalent bond.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_difference 1. .12 /note= "optionally D form residue"  
 FT  
 PN EP-752425-A2.  
 PD 08-JAN-1997.  
 PF 19-JUN-1996; 201706.  
 PR 21-JUN-1995; IT-MI1328.  
 PR 29-APR-1996; IT-MI0831.  
 PA (TECN-) TECNOGEN SCPA.  
 PI Fassina G, Ruvo M, Verdoliva A;  
 DR WPI; 97-067416/07.  
 DT New immunoglobulin-binding peptide(s) - for use in affinity  
 FT chromatography and immunoassay  
 PS Example 2; Page 7; 12pp; English.  
 CC W12305-W12306 are peptides that non-covalently bind to the constant  
 CC region of an immunoglobulin. They are used as affinity ligands for  
 CC separation and purification of immunoglobulins (Ig), and as capture  
 CC ligands for binding Ig to microtitre plates for ELISA determination.  
 SQ Sequence 15 AA;

Query Match 75.0%; Score 36; DB 1; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 0.23;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RYRVRYV 8  
 |||||  
 Db 3 RYTRY 10

RESULT 7  
 W89250  
 ID W89250 standard; Protein; 463 AA.  
 AC W89250;  
 DT 10-MAR-1999 (first entry)  
 DE Mouse PTP05 isoform #1.  
 KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;  
 KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;  
 KW neurodegenerative disease; neuronal survival; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease.  
 OS Mus sp.  
 PN W09849317-A2.  
 PD 05-NOV-1998.  
 PF 27-APR-1998; U08439.  
 PR 23-OCT-1997; US-063595.  
 PR 28-APR-1997; US-044428.  
 PR 20-MAY-1997; US-047222.  
 PR 11-JUN-1997; US-049477.  
 PR 11-JUN-1997; US-049756.  
 PR 18-JUN-1997; US-049914.  
 PA (SUG)- SUGEN INC.  
 PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,  
 PI Markby D, Onrust S, Peles E, Plozman GD;  
 DR WPI; 99-009434/01.  
 DT New nucleic acid encoding specific protein tyrosine phosphatases -  
 PT useful for identifying specific modulators for treatment and  
 PT prevention of cancer and neurodegenerative disease

Qy 1 RYRVRYV 9  
 |||||  
 Db 376 FIKYRVV 383

RESULT 8  
 W89251  
 ID W89251 standard; Protein; 405 AA.  
 AC W89251;  
 DT 10-MAR-1999 (first entry)  
 DE Mouse PTP05 isoform #2.  
 KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;  
 KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;  
 KW neurodegenerative disease; neuronal survival; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease.  
 OS Mus sp.  
 PN W09849317-A2.  
 PD 05-NOV-1998.  
 PF 27-APR-1998; U08439.  
 PR 23-OCT-1997; US-063595.  
 PR 28-APR-1997; US-044428.  
 PR 20-MAY-1997; US-047222.  
 PR 11-JUN-1997; US-049477.  
 PR 11-JUN-1997; US-049756.  
 PR 18-JUN-1997; US-049914.  
 PA (SUG)- SUGEN INC.  
 PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,  
 PI Markby D, Onrust S, Peles E, Plozman GD;  
 DR WPI; 99-009434/01.  
 DT New nucleic acid encoding specific protein tyrosine phosphatases -  
 PT useful for identifying specific modulators for treatment and  
 PT prevention of cancer and neurodegenerative disease

PS Claim 2; Page 157-158; 193pp; English.  
 CC The present invention describes isolated, enriched or purified nucleic  
 CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The  
 CC present sequence represents mouse PTP05. The above proteins, other than  
 CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify  
 CC substances that modulate their activity (i.e. agonists and antagonists,  
 CC including NBP) in vivo or in vitro. These substances are used to treat  
 CC or prevent diseases associated with abnormal signal transduction  
 CC pathways that involve the proteins, particularly cancer (e.g. leukaemia  
 CC and lymphoma), while modulators of ALK-7 (which is a type I receptor  
 CC serine/threonine kinase) are used to promote neuronal survival,  
 CC particularly for treating Alzheimer's, Parkinson's or Huntington's  
 CC diseases. Nucleic acid fragments of the polynucleotides encoding the  
 CC proteins can be used as probes to identify and clone related sequences;  
 CC to detect protein-encoded RNA; to generate transgenic animals and in  
 CC gene therapy (optionally after mutation). Ab are used to determine the  
 CC proteins. Sequence 463 AA;  
 SQ

Query Match 72.9%; Score 35; DB 1; Length 463;  
 Best Local Similarity 62.5%; Pred. No. 16;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YRVRYV 9  
 ::|||  
 Db 376 FIKYRVV 383

us-09-124-280a-39.rag

```

CC proteins.
SQ Sequence 405 AA;

Query Match 72.9%; Score 35; DB 1; Length 405;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRYVRV 9
DB 339 FIKYVRV 346

RESULT 9
ID W89252 standard; Protein; 122 AA.
AC W89252;
DE Rat PTP10.
KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
KW neurodegenerative disease; neuronal survival; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease.
OS Rattus sp.
PN WO9849317-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08439.
PR 23-OCT-1997; US-063595.
PR 28-APR-1997; US-044428.
PR 20-MAY-1997; US-047222.
PR 11-JUN-1997; US-049477.
PR 11-JUN-1997; US-049756.
PR 18-JUN-1997; US-049914.
PA (SUGEN-) SUGEN INC.
PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,
PI Markby D, Onrust S, Peles E, Plowman GD;
DR WPI: 99-009434/01.
DR N-PSDB; V81747.
PT New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
PS Claim 2; Page 160; 193pp; English.
CC The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence represents rat PTP10. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC or prevent diseases associated with abnormal signal transduction
CC including NBP) in vivo or in vitro. These substances are used to treat
CC pathways that involve the proteins, particularly cancer (e.g. leukaemia
CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
SQ Sequence 122 AA;

Query Match 72.9%; Score 35; DB 1; Length 122;
Best Local Similarity 62.5%; Pred. No. 3.6;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRYVRV 9
DB 96 FIKYVRV 103

RESULT 10
W89249
ID W89249 standard; Protein; 426 AA.

```

```

AC W89249;
DE 10-MAR-1999 (first entry)
DE Mouse PTP05.
KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
KW type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
KW neurodegenerative disease; neuronal survival; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease.
OS Mus sp.
PN WO9849317-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08439.
PR 23-OCT-1997; US-063595.
PR 28-APR-1997; US-044428.
PR 20-MAY-1997; US-047222.
PR 11-JUN-1997; US-049477.
PR 11-JUN-1997; US-049756.
PR 18-JUN-1997; US-049914.
PA (SUGEN-) SUGEN INC.
PI App H, Clary D, Courtneidge SA, Hui TH, Jallal B,
PI Markby D, Onrust S, Peles E, Plowman GD;
DR WPI: 99-009434/01.
DR N-PSDB; V81744.
PT New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
PS Claim 2; Page 155-157; 193pp; English.
CC The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence represents mouse PTP05. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC including NBP) in vivo or in vitro. These substances are used to treat
CC or prevent diseases associated with abnormal signal transduction
CC pathways that involve the proteins, particularly cancer (e.g. leukaemia
CC and lymphoma), while modulators of ALK-7 (which is a type I receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
SQ Sequence 426 AA;

Query Match 72.9%; Score 35; DB 1; Length 426;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRYVRV 9
DB 339 FIKYVRV 346

RESULT 11
W88313
ID W88313 standard; Protein; 394 AA.
AC W88313;
DE 26-APR-1999 (first entry)
DE O antigen polymerase involved in O157 antigen biosynthesis.
KW O antigen; O157 antigen; polymerase; wzy gene; diarrhoea;
KW haemorrhagic colitis; diagnosis.
OS Escherichia coli.
PN WO9850531-A1.
PD 12-NOV-1998.
PF 01-MAY-1998; AU0315.
PR 22-JUL-1997; AU-008162.
PR 01-MAY-1997; AU-006545.
PA (UNSY ) UNIV SYDNEY.
PI Reeves PR, Wang L;
DR WPI: 99-059669/05.
DR N-PSDB; X06749.
PT Nucleic acid molecules specific for bacterial polysaccharide

```

PT antigens - useful for detecting specific strains in, e.g. food,  
 PT faeces or patient samples  
 PS Disclosure; Fig 8; 165pp; English.  
 CC This is the amino acid sequence of an O antigen polymerase that is  
 CC encoded by the wzy open reading frame of a gene cluster (see  
 CC X06749) involved in the biosynthesis of the Escherichia coli O157  
 CC O antigen. The use of nucleic acid molecules derived from  
 CC particular assembly and transport genes, particularly wbd  
 CC (transferase), wzx (flippase) and wzy (polymerase) genes, within O  
 CC antigen gene clusters improves the specificity of methods for the  
 CC detection and identification of O antigens, e.g. in testing food-  
 CC or faecal-derived samples, or samples from patients. The O antigen  
 CC is a major virulence factor of enteropathogenic E. coli strains  
 CC that cause diarrhoea and haemorrhagic colitis.  
 SQ Sequence 394 AA;

Query Match 70.8%; Score 34; DB 1; Length 394;  
 Best Local Similarity 57.1%; Pred. No. 20;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 YVRYVRY 8  
 I::I::I  
 Db 117 YIRYIKY 123

## RESULT 12

R37386  
 ID R37386 standard; peptide; 6 AA.  
 AC R37386;

DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B; Group II.  
 OS Synthetic.

EH Key Location/Qualifiers  
 FT region 1..3  
 FT /note= "repeat region"  
 FT region 4..6  
 FT /note= "repeat region"

PN ZAS200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI Porro M;

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated

PS Example; Page 5; 39pp; English.

CC The (Group II) peptide is an example of a generic peptide of formula  
 CC R-(Lys/Arg/His-Phe/Tyr/Trp-Leu/Ile/Val)n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock,

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

SQ Sequence 6 AA;

Query Match 66.7%; Score 32; DB 1; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 RYRYRV 6  
 I::I::I  
 Db 1 RYRYRV 6

## RESULT 13

R82070  
 ID R82070 standard; Protein; 1422 AA.  
 AC R82070;

DT 02-JUL-1996 (first entry)  
 DE Hepatitis GB virus (HGBV) clone protein prod.  
 DE Hepatitis GB virus; HGBV; diagnosis; treatment; vaccine;  
 KW reagents; non-A; non-B; non-C; non-D; non-E; clone;  
 KW tamarin; infected plasma; lambda phage; cDNA library.  
 OS Hepatitis GB virus.

EH Key Location/Qualifiers

FT misc\_difference 1..1422  
 FT /note= "others correspond to degenerate or STOP  
 FT codons in T00052"

PN WO9521922-A2.

PD 17-AUG-1995.

PF 14-FEB-1995; U02118.

PR 14-FEB-1994; US-196030.

PR 13-MAY-1994; US-242654.

PR 29-JUL-1994; US-283314.

PR 23-NOV-1994; US-344190.

PR 23-NOV-1994; US-344185.

PR 27-JAN-1995; US-344557.

PA (ABBO) ABBOTT LAB.

PI Buijk SL, Dawson GJ, Desai SM, Erker JC, Leary TP;

PI Muerhoff AS, Mushahwar IK, Pilot-Matias TJ, Schlauder GG;

PI Simons JN;

DR WPI; 95-293123/38.

DR N-PSDB; T00052.

PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful  
 PT for diagnosis and therapy of hepatitis GB virus

PS Example 5; Pages 254-259; 661pp; English.

CC Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV

CC infected tamarin plasma, using standard procedures, was used to

CC prepare a lambda phage HGBV cDNA library. The cDNA clone T00052,

CC which encodes the proteins R82066-71 (the 6 possible reading

CC frames), was rescued from the lambda phage, searched against a

CC sequence database and found to be an unique HGBV sequence.

CC Reagents which comprise the HGBV DNA, or its protein prods. can

CC be used for the diagnosis, therapy or in a vaccine to prevent

CC HGBV infection.

SQ Sequence 1422 AA;

Query Match 64.6%; Score 31; DB 1; Length 1422;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 RYRYRVYV 9  
 I::I::I  
 Db 1213 RYMPVYLV 1221

## RESULT 14

R13944  
 ID R13944 standard; Protein; 553 AA.  
 AC R13944;

DT 27-NOV-1991 (first entry)

DE Partial HVT ribonucleotide reductase large subunit RRI.

KW Newcastle disease; recombinant poultry virus vaccine.

OS Turkey herpes virus strain FC 126.

PN EP-447303-A.

PD 18-SEP-1991.

PF 07-MAR-1991; 400634.

PR 12-MAR-1990; FR-003105.

PA (RHON-) RHONE MERIEUX SA.

PI Rey-Senelonge A, Kohen G;  
DR WPI; 91-275886/38.  
DR N-PSDB; Q13430.

PT New herpes recombinants - useful as vaccines against human and  
PT animal viral conditions, e.g. fowl pest, coccidiosis and  
PT pasteurellosis

PS Disclosure: Page 13-14; 22pp; French.

CC This sequence is the C-terminal region of HVT RRL. It was deduced  
CC from a genomic DNA clone which was isolated because it contained  
CC the entire RR2 gene. Knowledge of the RR2 small sub-unit sequence is  
CC exploited in the construction of recombinant vaccines. A heterologous  
CC sequence coding for an appropriate immunogen to protect against e.g.  
CC Newcastle disease, avian anaemia, colibacillosis, avian infectious  
CC bronchitis, etc. is inserted into the RR2 gene. The RVT is suitable  
CC for use in live vaccines as it is apathogenic and non-oncogenic.  
CC See also Q13668.

SQ Sequence 553 AA;

Query Match 62.5%; Score 30; DB 1; Length 553;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YVRYVRYV 9

Db 108 FRYLYRYV 115

#### RESULT 15

R37387  
ID R37387 standard; peptide; 6 AA.

AC R37387;

DT 07-JUL-1993 (first entry)

DE Peptide for treating septic shock.

KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;

KW cytokine release control; treatment; pertussis; bacterial meningitis;

KW HIV related infections; polymyxin B; Group II.

OS Synthetic.

FH Key Location/Qualifiers

FT region 1..3

FT /note= "repeat region"

FT region 4..6

FT /note= "repeat region"

FT ZA9200943-A.

PD 25-NOV-1992.

PF 10-FEB-1992; 000943.

PR 11-FEB-1991; US-658744.

PA (PORR/) PORRO M.

PI PORRO M;

DR WPI; 93-094304/11.

PT New peptide for treatment or prevention of toxic shock - comprises

PT specified sequences of aminoacid(s) and analogs

PT comprising sequences retro-orientated

PS Example: Page 5; 39pp; English.

CC The (Group II) peptide is an example of a generic peptide of formula

CC R-(Lys/Arg/His - Phe/Tyr/Trip - Leu/Ile/Val)n-R, where n = 1-100

CC and each R is H, an amino acid residue or a fatty acid residue.

CC The peptide is useful for treating or preventing septic shock,

CC mixing with polymyxin B to reduce its toxicity; removing

CC endotoxins from blood, sera or other fluids (in vivo or in

CC vitro); controlling release of cytokines induced by endotoxins;

CC as diagnostic reagents to detect and quantify toxins in blood

CC or sera; preparing non-toxic antigenic complexes of lipid A or

CC lipopolysaccharide (LPS); and for treating pertussis, bacterial

CC meningitis and HIV-related infections. The usual dose is 10-100

CC ug/kg/day, given parenterally. It binds to the same sites as

CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It

CC has no antibiotic activity; does not lyse erythrocytes; has no

CC toxicity in mice when injected at 50mg/kg and is relatively unstable

CC against proteases.

SQ Sequence 6 AA;

Query Match 62.5%; Score 30; DB 1; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.5e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYRYV 6

Db 1 RYIRYI 6

Search completed: September 7, 1999, 22:50:09  
Job time: 7890 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.  
OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:51 ; Search time 82.37 Seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-39  
Perfect score: 48  
Sequence: 1 RYRVRYRV 9  
Scoring table: BLOSUM62  
Searched: 106577 seqs, 9868381 residues  
Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	9	1	US-07-819-893-10
2	48	100.0	9	1	US-08-280-397-10
3	48	100.0	9	1	US-08-097-830E-9
4	48	100.0	9	2	US-08-456-112B-9
5	48	100.0	9	2	US-08-456-112B-39
6	33	68.8	9	1	US-08-097-830E-31
7	30	62.5	990	1	US-08-232-540-2
8	30	62.5	1013	1	US-08-233-008A-8
9	30	62.5	990	1	US-08-428-949A-2
10	30	62.5	571	1	US-08-368-803-17
11	30	62.5	990	1	US-08-428-948A-2
12	30	62.5	990	2	US-08-428-946-2
13	30	62.5	200	2	US-08-606-143-42
14	30	62.5	990	3	PCT-US95-04656-2
15	29	60.4	554	1	US-07-839-433-2
16	29	60.4	377	1	US-08-454-097-31
17	29	60.4	347	1	US-08-454-097-33
18	29	60.4	448	2	US-08-231-342-23
19	28	58.3	287	1	US-08-457-245-3
20	27	56.2	577	1	US-07-820-154A-30
21	27	56.2	1167	1	US-08-485-568A-6
22	27	56.2	1168	1	US-08-620-717A-9
23	27	56.2	599	2	US-08-752-238-3
24	27	56.2	1167	2	US-08-590-554A-6
25	27	56.2	536	2	US-08-551-211-3
26	27	56.2	577	2	US-08-663-566A-11
27	27	56.2	577	2	US-08-097-554A-30
28	27	56.2	2254	2	US-08-286-819A-28
29	27	56.2	788	2	US-08-918-914-4
30	27	56.2	457	2	US-08-882-704A-6
31	27	56.2	577	3	PCT-US93-00324-30
32	27	56.2	140	3	PCT-US93-11612-4
33	27	56.2	121	3	PCT-US93-11612-7
34	27	56.2	121	3	PCT-US93-11612-8
35	27	56.2	140	3	PCT-US93-11612-12
36	27	56.2	581	3	PCT-US94-01826A-13
37	27	56.2	581	3	PCT-US94-02252A-13
38	27	56.2	577	3	PCT-US95-10245-11
39	27	56.2	78	4	5432081-9

Sequence 3, Appl1  
Sequence 3, Appl1  
Sequence 30, Appl1  
Sequence 23, Appl1  
Sequence 112, App  
Sequence 14, Appl1

ALIGNMENTS

RESULT 1  
US-07-819-893-10  
; Sequence 10, Application US/07819893  
; Patent No. 5371186  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/819,893  
; FILING DATE: 19920115  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: AMINO ACIDS  
; TOPOLOGY: linear  
US-07-819-893-10

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

Qy 1 RYRVRYRV 9  
Db 1 RYRVRYRV 9

RESULT 2  
US-08-280-397-10  
; Sequence 10, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/26/94  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acids  
;; TOPOLOGY: linear  
;; US-08-280-397-10

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
Db 1 RYRVRYRV 9

RESULT 3  
US-08-097-830E-9  
; Sequence 9, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:

;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; US-08-097-830E-9

Query Match 100.0%; Score 48; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
Db 1 RYRVRYRV 9

RESULT 4  
US-08-456-112B-9  
; Sequence 9, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-456-112B-9

Query Match 100.0%; Score 48; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYRV 9  
Db 1 RYRVRYRV 9



Db 1 RYRVRYV 9

## RESULT 5

US-08-456-112B-39

; Sequence 39, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson & Costigan

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: LEADING EDGE 486

; OPERATING SYSTEM: DOS

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,112B

; FILING DATE: May 31, 1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Costigan, James V.

; REGISTRATION NUMBER: 25,669

; REFERENCE/DOCKET NUMBER: 576-004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 302-8998

; TELEFAX: (212) 302-8998

; INFORMATION FOR SEQ ID NO: 39:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids

; TYPE: amino acid

; TOPOLOGY: circular

; US-08-456-112B-39

Query Match

Best Local Similarity 100.0%; Score 48; DB 2; Length 9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYRVRYV 9

|||||

Db 1 RYRVRYV 9

## RESULT 6

US-08-097-830E-31

; Sequence 31, Application US/08097830E

; Patent No. 5652211

; GENERAL INFORMATION:

; APPLICANT: Porro, Massimo

; TITLE OF INVENTION: Peptides For Neutralizing The

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hedman, Gibson & Costigan, P.C.

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8998  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; US-08-097-830E-31

Query Match 68.8%; Score 33; DB 1; Length 9;

Best Local Similarity 75.0%; Pred. NO. 7.9e+04;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRVRYV 9

|||

Db 2 YVEYVYV 9

## RESULT 7

US-08-232-540-2

; Sequence 2, Application US/08232540

; Patent No. 5498536

; GENERAL INFORMATION:

; APPLICANT: Khandke, Kiran M.

; TITLE OF INVENTION: A No. 5498536el Protein Designated

; TITLE OF INVENTION: Chondroitinase II and its Use With A Protein Designated

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: American Cyanamid Company

; STREET: One Cyanamid Plaza

; CITY: Wayne

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07470-8426

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,540

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Gordon, Alan M.

; REGISTRATION NUMBER: 30,637

; REFERENCE/DOCKET NUMBER: 32,390

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-831-3244

; TELEFAX: 201-831-3305

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 990 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-232-540-2

Query Match 62.5%; Score 30; DB 1; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRYVRY 8  
|||::|  
Db 615 RYGRYLOY 622

RESULT 8  
US-08-233-008A-8  
; Sequence 8, Application US/08233008A  
; Patent No. 5578480  
; GENERAL INFORMATION:  
; APPLICANT: Khandke, Kiran M.  
; TITLE OF INVENTION: Methods For The Isolation And  
; TITLE OF INVENTION: Purification Of The Recombinantly Expressed chondroitinase  
; TITLE OF INVENTION: I and II Enzymes From P. vulgaris  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company  
; STREET: One Cyanamid Plaza  
; CITY: Wayne  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07470-8426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/233.008A  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gordon, Alan M.  
; REGISTRATION NUMBER: 30,637  
; REFERENCE/DOCKET NUMBER: 31,885-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-831-3244  
; TELEFAX: 201-831-3305  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-233-008A-8

Query Match 62.5%; Score 30; DB 1; Length 1013;  
Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRYVRY 8  
|||::|  
Db 638 RYGRYLOY 645

RESULT 9  
US-08-428-949A-2  
; Sequence 2, Application US/08428949A  
; Patent No. 5716617  
; GENERAL INFORMATION:  
; APPLICANT: Kiran M. Khandke, John Gotto, Ursula Eul  
; TITLE OF INVENTION: Compositions of Chondroitinase I and  
; TITLE OF INVENTION: Chondroitinase II  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby P.C.

STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,949A  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Robinson, Joseph R.  
; REGISTRATION NUMBER: 33,448  
; REFERENCE/DOCKET NUMBER: 0646/1B017-US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-527-7783  
; TELEFAX: 212-753-6237  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 990 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-949A-2

Query Match 62.5%; Score 30; DB 1; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRYVRY 8  
|||::|  
Db 615 RYGRYLOY 622

RESULT 10  
US-08-368-803-17  
; Sequence 17, Application US/08368803  
; Patent No. 5733554  
; GENERAL INFORMATION:  
; APPLICANT: AUDONNET, Jean-Christophe F  
; APPLICANT: BUBLOT, Michel J  
; APPLICANT: DARTEIL, Raphael J  
; APPLICANT: DUINAT, Carole V  
; APPLICANT: LAPLACE, Eliane L  
; APPLICANT: RIVIERE, Michel A  
; TITLE OF INVENTION: Avian Herpesvirus-based live recombinant avian  
; TITLE OF INVENTION: vaccine, in particular against Gumboro disease  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LARSON AND TAYLOR  
; STREET: 727 SOUTH 23RD STREET  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/368,803  
; FILING DATE: 05-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SARRO, Thomas  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 920-7200

TELEFAX: (703) 892-8428  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 571 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-368-803-17

Query Match 62.5%; Score 30; DB 1; Length 571;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRYRYRY 8  
||| |||  
Db 333 RYIYKRY 340

RESULT 11  
US-08-428-948A-2  
Sequence 2, Application US/08428948A  
Patent No. 5741692  
GENERAL INFORMATION:  
APPLICANT: Khandke, Kiran M.  
TITLE OF INVENTION: Chondroitinase II  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,948A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Robinson, Joseph R.  
REGISTRATION NUMBER: 33,448  
REFERENCE/DOCKET NUMBER: 0646/1B017-US4  
TELEPHONE: 212-527-7783  
TELEFAX: 212-753-6327  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-948A-2

Query Match 62.5%; Score 30; DB 1; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRYRYRY 8  
||| |||  
Db 615 RYGRLOY 622

RESULT 12  
US-08-428-946-2  
Sequence 2, Application US/08428946  
Patent No. 5855883  
GENERAL INFORMATION:

APPLICANT: Kiran M. Khandke, John Gatto, Ursula Eul  
TITLE OF INVENTION: Method of Disinsection of Vitreous Body from  
Neural Retina of the Eye  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby & Darby P.C.  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,946  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Robinson, Joseph R.  
REGISTRATION NUMBER: 33,448  
REFERENCE/DOCKET NUMBER: 0646/1B017-US3  
TELEPHONE: 212-527-7783  
TELEFAX: 212-753-6237  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-946-2

Query Match 62.5%; Score 30; DB 2; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRYRYRY 8  
||| |||  
Db 615 RYGRLOY 622

RESULT 13  
US-08-606-143-42  
Sequence 42, Application US/08606143  
Patent No. 5856155  
GENERAL INFORMATION:  
APPLICANT: Li, Min  
TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR  
MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH  
TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH  
TITLE OF INVENTION: COMPOUNDS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Volt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,143  
FILING DATE: 23-FEB-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kilyk Jr., John

REGISTRATION NUMBER: 30763  
REFERENCE/DOCKET NUMBER: 71756  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 616-5600  
TELEFAX: (312) 616-5700  
TELEX: 25-3533  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 200 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-606-143-42

Query Match 62.5%; Score 30; DB 2; Length 200;  
Best Local Similarity 62.5%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RYRVRY 8  
|||:  
Db 181 RYRVAF 188

RESULT 14  
PCT-US95-04656-2  
Sequence 2, Application PC/TUS9504656  
GENERAL INFORMATION:  
APPLICANT: Khandke, Kiran M.  
TITLE OF INVENTION: A Novel Protein Designated  
TITLE OF INVENTION: Chondroitinase II and its Use With a Protein Designated  
TITLE OF INVENTION: Chondroitinase I to Achieve Complete Vitreal Disinsertion  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04656  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gordon, Alan M.  
REGISTRATION NUMBER: 30,637  
REFERENCE/DOCKET NUMBER: 32,390-00/PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-831-3244  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-04656-2

Query Match 62.5%; Score 30; DB 3; Length 990;  
Best Local Similarity 62.5%; Pred. No. 2,4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVRY 8  
|||:  
Db 615 RYRQLQY 622

RESULT 15  
US-07-839-433-2  
Sequence 2, Application US/07839433  
Patent No. 5420021  
GENERAL INFORMATION:  
APPLICANT: MARUGG, JOHN D  
APPLICANT: TOONEN, MARIA Y  
APPLICANT: VERHUE, WALTER M  
APPLICANT: VERRIPS, CORNELIS T  
TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF  
TITLE OF INVENTION: ALPHA-ACETOLACTIC ACID  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DABBY & CUSHMAN  
STREET: 1615 L STREET, N.W.  
CITY: WASHINGTON, D.C.  
COUNTRY: U.S.A.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Tape  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/839,433  
FILING DATE: 19920224  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KORULIS, PAUL N  
REGISTRATION NUMBER: 16773  
REFERENCE/DOCKET NUMBER: 94713  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 554 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-839-433-2

Query Match 60.4%; Score 29; DB 1; Length 554;  
Best Local Similarity 71.4%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YRVRY 8  
|||:  
Db 493 YVDVYKY 499

Search completed: September 7, 1999, 23:07:51  
Job time: 1749 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:57 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-39  
Perfect score: 48  
Sequence: 1 RYRVRYV 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	72.9	110	1 S73233	ribosomal protein
2	34	70.8	261	2 H69217	hypothetical prote
3	34	70.8	675	2 D48492	kpsc protein - Esc
4	34	70.8	600	2 S48509	transcription regu
5	33	68.8	1294	2 S77690	probable membrane
6	33	68.8	285	2 C65012	hypothetical prote
7	33	68.8	113	2 F64001	hypothetical prote
8	33	68.8	630	2 C39390	hypothetical prote
9	32	66.7	129	2 S53699	ribosomal protein
10	32	66.7	638	2 G02068	white homolog - hu
11	32	66.7	322	2 A70189	hypothetical prote
12	32	66.7	400	2 S32879	probable membrane
13	32	66.7	481	2 H70879	lipa protein - Nei
14	32	66.7	208	2 C64347	hypothetical prote
15	32	66.7	502	2 S26004	18S rRNA intron 1
16	32	66.7	785	2 S63652	hypothetical prote
17	32	66.7	507	2 S52648	myo-inositol-1-pho
18	31	64.6	813	1 A49123	fibroblast growth
19	31	64.6	331	2 S69192	serine o-acetyltra
20	31	64.6	480	2 A56182	fibroblast growth
21	31	64.6	480	2 B56182	fibroblast growth
22	31	64.6	170	2 D65001	hypothetical prote
23	31	64.6	165	2 C42465	beta protein homol
24	31	64.6	365	2 E70791	probable lyase - M
25	31	64.6	1259	2 S25954	gene atpA intron 2
26	31	64.6	299	2 S35315	cdc16 protein - fi
27	31	64.6	514	2 C70446	hypothetical prote
28	31	64.6	223	2 B71057	hypothetical prote
29	30	62.5	1603	1 BVASAL	arom protein - Eme
30	30	62.5	570	1 HNNZAV	hemagglutinin-neur
31	30	62.5	571	1 H46328	hemagglutinin-neur
32	30	62.5	571	1 I46328	hemagglutinin-neur
33	30	62.5	571	1 A36829	hemagglutinin-neur
34	30	62.5	571	1 B36829	hemagglutinin-neur
35	30	62.5	571	2 S07126	hemagglutinin-neur
36	30	62.5	571	2 S40164	hemagglutinin-neur
37	30	62.5	581	2 S45114	hemagglutinin-neur
38	30	62.5	436	2 S66171	ribulose-bisphosph
39	30	62.5	228	2 C69859	two-component resp

peptide transport  
transcription regu  
probable membrane  
potassium channel  
potassium channel  
delayed rectifier

ALIGNMENTS

RESULT 1

S73233

ribosomal protein L23 - red alga (Porphyra purpurea) chloroplast

C:Species: chloroplast Porphyra purpurea

C>Date: 23-Oct-1998 #sequence\_revision 23-Oct-1998 #text\_change 23-Oct-1998

C:Accession: S73233

R:Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A:Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A:Reference number: S73108

A:Accession: S73233

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-110 <REI>

A:Cross-references: EMBL:U38804; NID:g1276652; PID:g1276778

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: rpl23

A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein L23

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 72.9%; Score 35; DB 1; Length 110;

Best Local Similarity 55.8%; Pred. No. 4.6;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVRYV 9

DB 2 RYLKRYII 10

RESULT 2

H69217

hypothetical protein MTH882 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Sep-1998

C:Accession: H69217

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

; Olu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,

ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: H69217

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <MTH>

A:Cross-references: GB:AE000864; GB:AE000666; NID:g2621970; PID:g2621976

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH882

A:Start codon: TTG

C:Superfamily: hypothetical protein HI0902

Query Match 70.8%; Score 34; DB 2; Length 261;

Best Local Similarity 71.4%; Pred. No. 17;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYRVRYV 9

DB 3 RYLKRYII 10

Db 241 VRYIRYI 247

RESULT 3

D48492

kpsc protein - Escherichia coli

C:Species: Escherichia coli

C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 09-Sep-1997

C:Accession: D48492; I67694; S36652

R:Pazzani, C.; Rosenow, C.; Boulnois, G.J.; Bronner, D.; Jann, K.; Roberts, I.S.

J. Bacteriol. 175, 5978-5983, 1993

A:Title: Molecular analysis of region 1 of the Escherichia coli K5 antigen gene cluster:

A:Reference number: 448492; MUID:93388530

A:Accession: D48492

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-675 <PAZ>

A:Cross-references: EMBL:X74567; NID:G397404; PID:G397408

R:Rosenow, C.; Roberts, I.S.; Jann, K.

FEMS Microbiol. Lett. 125, 159-164, 1995

A:Title: Isolation from recombinant Escherichia coli and characterization of CMP-Kdo synthetase from Escherichia coli

A:Reference number: 153591; MUID:95180691

A:Accession: I67694

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-8 <RES>

A:Cross-references: GB:S76943; NID:9913365

Query Match 70.8%; Score 34; DB 2; Length 675;

Best Local Similarity 62.5%; Pred. No. 43;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRYV 9

I: I I I I

Db 293 YLRYCRYI 300

RESULT 4

S48509

transcription regulator SPO8 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein G1337; protein YGL192W

C:Species: Saccharomyces cerevisiae

C:Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 06-Feb-1998

C:Accession: S48509; S62053; S64209

R:Sakurai, M.

submitted to the EMBL Data Library, November 1993

A:Reference number: S48509

A:Accession: S48509

A:Molecule type: DNA

A:Residues: 1-600 <SAK>

A:Cross-references: EMBL:D23721; NID:G433221; PID:dl005476; PID:g471168

R:Coglievina, M.; Delneri, D.; Zaccaria, P.; Klima, R.; Bertani, I.; Bruschi, C.V.

submitted to the EMBL Data Library, September 1995

A:Description: A 6.7 kb fragment from chromosome VII of Saccharomyces cerevisiae contains the SPO8 gene

A:Reference number: S62051

A:Accession: S62053

A:Molecule type: DNA

A:Residues: 1-600 <COG>

A:Cross-references: EMBL:X91837; NID:g1177627; PID:e203623; PID:g1177636

A:Experimental source: strain FY1679

R:Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64183

A:Accession: S64209

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <BRU>

A:Cross-references: EMBL:Z72714; NID:g1322814; PID:e243795; PID:g1322815; MIPS:YGL192W

C:Genetics:

A:Gene: SPO8

A:Cross-references: SGD:S0003160; MIPS:YGL192W

A:Map position: 7L

C:Keywords: nucleus; transcription regulation

Query Match 70.8%; Score 34; DB 2; Length 600;

Best Local Similarity 44.4%; Pred. No. 39;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RXRVYRVYV 9

I: I: I: I: I:

Db 272 RYHYLYYI 280

RESULT 5

S77690

probable membrane protein YOL075c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein Oll25; hypothetical protein Oll130; hypothetical

C:Species: Saccharomyces cerevisiae

C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Jul-1998

C:Accession: S77690; S66767; S66768

R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66756

A:Accession: S77690

A:Molecule type: DNA

A:Residues: 1-1294 <ALE>

A:Cross-references: EMBL:Z74816; MIPS:YOL075c

A:Note: this is a revision to the sequence from reference S66756

R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66756

A:Accession: S66767

A:Molecule type: DNA

A:Residues: 1-179, 'TTRTGVLVVKRED' <ALW>

A:Cross-references: EMBL:Z74816

A:Experimental source: strain S288C

A:Note: this sequence has been revised in reference S77690

A:Note: this was assumed to be protein YOL074c

A:Accession: S66768

A:Molecule type: DNA

A:Residues: 200-1294 <ALF>

A:Cross-references: EMBL:Z74817

A:Experimental source: strain S288C

A:Note: this sequence has been revised in reference S77690

A:Note: this was assumed to be the complete sequence of protein YOL075c

C:Genetics:

A:Map position: 15L

A:Note: YOL075c

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

C:Keywords: P-loop; transmembrane protein

F:45-263/Domain: ATP-binding cassette homology <ABC1>

F:62-69/Region: nucleotide-binding motif A (P-loop)

F:376-392/Domain: transmembrane #status predicted <TM1>

F:469-485/Domain: transmembrane #status predicted <TM2>

F:496-512/Domain: transmembrane #status predicted <TM3>

F:606-622/Domain: transmembrane #status predicted <TM4>

F:710-916/Domain: ATP-binding cassette homology <ABC2>

F:727-734/Region: nucleotide-binding motif A (P-loop)

F:1042-1058/Domain: transmembrane #status predicted <TM5>

F:1125-1141/Domain: transmembrane #status predicted <TM6>

F:1177-1193/Domain: transmembrane #status predicted <TM7>

F:1269-1285/Domain: transmembrane #status predicted <TM8>

Query Match 68.8%; Score 33; DB 2; Length 1294;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVRYVRYV 9

I: I: I: I: I:

Db 550 YVRWIKYI 557

RESULT 6

## C65012

hypothetical protein b2382 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 18-Sep-1998  
C:Accession: C65012  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: C65012  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-285 <BLAT>  
A:Cross-references: GB:AF000326; GB:U00096; NID:gl1788718; PID:gl1788725; UWGP:b2382  
A:Experimental source: strain K-12, substrain M61655  
C:Superfamily: hypothetical protein b2382

Query Match 68.8%; Score 33; DB 2; Length 285;  
Best Local Similarity 55.6%; Pred. No. 28;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVYRVV 9

| : : | | | : |

Db 225 RFIEYRVW 233

## RESULT 7

F64001  
hypothetical protein HI0109 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 10-Oct-1997  
C:Accession: F64001  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: F64001  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-113 <TIGR>  
A:Cross-references: GB:U32696; GB:L42023; NID:gl573057; PID:gl573066; TIGR:HI0109

Query Match 68.8%; Score 33; DB 2; Length 113;  
Best Local Similarity 62.5%; Pred. No. 11;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YRVYRVV 9

: | : | | |

Db 79 FVEYRVV 86

## RESULT 8

C39930  
hypothetical protein (internalin region) - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 18-Jun-1993  
C:Accession: C39930  
R:Gallard, J.L.; Berche, P.; Frehel, C.; Gouin, E.; Cossart, P.  
Cell 65, 1127-1141, 1991  
A:Title: Entry of Listeria monocytogenes into cells is mediated by internalin, a repeat  
A:Reference number: A39930; MUID:91292517  
A:Accession: C39930  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-630 <GAI>  
A:Cross-references: GB:M67471

Query Match 68.8%; Score 33; DB 2; Length 630;  
Best Local Similarity 71.4%; Pred. No. 62;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYRVYRV 7

| : : | | | : |

Db 401 RYVYIR 407

## RESULT 9

S53699  
ribosomal protein SL31 - Sulfolobus acidocaldarius  
C:Species: Sulfolobus acidocaldarius  
C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S53699; S41964  
R:Moll, R.; Schmidtko, S.; Schaefer, G.  
Biochim. Biophys. Acta 1261, 315-318, 1995  
A:Title: Nucleotide sequence of a gene cluster encoding ribosomal proteins in the the  
A:Reference number: S53698; MUID:95226466  
A:Accession: S53699  
A:Molecule type: DNA  
A:Residues: 1-129 <MOL>  
A:Cross-references: EMBL:X77509; NID:g453437; PID:g300459; PID:g453439  
C:Superfamily: rat ribosomal protein L31  
C:Keywords: protein biosynthesis; ribosome

Query Match 66.7%; Score 32; DB 2; Length 129;  
Best Local Similarity 55.6%; Pred. No. 19;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVYRVV 9

| : : | | | : |

Db 68 RAIYVRYM 76

## RESULT 10

G02068  
white homolog - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 28-Aug-1998  
C:Accession: G02068  
R:Croop, J.M.; Tiller, G.; Fletcher, J.A.; Lux, M.; Raab, E.; Goldenson, D.; Arcinieg  
submitted to the EMBL Data Library, August 1995  
A:Reference number: H00769  
A:Accession: G02068  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-638 <CRO>  
A:Cross-references: EMBL:U34919; NID:gl314276; PID:gl314277  
C:Genetics:  
A:Gene: white  
C:Superfamily: fruit fly white protein; ATP-binding cassette homology  
C:Keywords: P-loop  
F:61-253/Domain: ATP-binding cassette homology <ABC>  
F:78-85/Region: nucleotide-binding motif A (P-loop)

Query Match 66.7%; Score 32; DB 2; Length 638;  
Best Local Similarity 71.4%; Pred. No. 95;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YRVYRVV 8

| : | | | |

Db 558 YISYRVY 564

## RESULT 11

A70189  
hypothetical protein BB0714 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998

C:Accession: A70189  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, B.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: A70189

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-322 <KLE>

A:Cross-references: GB:AB001171; GB:AE000783; NID:g2688640; PID:g2688650; TIGR:BB0714

A:Experimental source: strain B31

C:Superfamily: hypothetical protein BB0714; tetratricopeptide repeat homology

F:122-135/Domain: tetratricopeptide repeat homology <Tr1>

F:161-193/Domain: tetratricopeptide repeat homology #status atypical <WT2>

F:194-227/Domain: tetratricopeptide repeat homology <Tr3>

F:240-273/Domain: tetratricopeptide repeat homology <Tr4>

Query Match 66.7%; Score 32; DB 2; Length 322;

Best Local Similarity 62.5%; Pred. No. 48;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVRYV 8

Db 4 RYLRIFY 11

RESULT 12

lipA protein - Neisseria meningitidis

C:Species: Neisseria meningitidis

C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 03-Aug-1995

C:Accession: S32879; S28077

R:Frosch, M.; Mueller, A.

Mol. Microbiol. 8, 483-493, 1993

A:Title: Phospholipid substitution of capsular polysaccharides and mechanisms of capsule

A:Reference number: S32879

A:Accession: S32879

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <FRO>

A:Cross-references: EMBL:Z13995

C:Genetics:

A:Gene: lipA

Query Match 66.7%; Score 32; DB 2; Length 400;

Best Local Similarity 55.6%; Pred. No. 59;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVRYV 9

Db 134 RYLYQSYRL 142

RESULT 13

H70679

probable membrane protein with some - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998

C:Accession: H70679

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: H70679

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-481 <COL>

A:Cross-references: GB:Z81451; GB:AL123456; NID:g3261662; PID:g280499; PID:g1666150

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV2434c

Query Match 66.7%; Score 32; DB 2; Length 481;

Best Local Similarity 75.0%; Pred. No. 71;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVRYV 8

Db 348 RYARLVRY 355

RESULT 14

C64347

hypothetical protein MJ0379 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 10-Oct-1997

C:Accession: C64347

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc

A:Reference number: A64300; MUID:96337999

A:Accession: C64347

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-208 <BUL>

A:Cross-references: GB:U67490; GB:L77117; NID:g1591074; PID:g1591084; TIGR:MJ0379; PI

C:Genetics:

A:Map position: REV345512-344886

A:Start codon: TTG

Query Match 66.7%; Score 32; DB 2; Length 208;

Best Local Similarity 71.4%; Pred. No. 31;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RYRVRYV 9

Db 2 RYRMRYI 8

RESULT 15

S26004

18S rRNA intron 1 protein - liverwort (Marchantia polymorpha) mitochondrion

C:Species: mitochondrion Marchantia polymorpha

C>Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 17-Mar-1999

C:Accession: S26004

R:Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.;

J. Mol. Biol. 223, 1-7, 1992

A:Title: Gene organization deduced from the complete sequence of liverwort Marchantia

A:Reference number: S25941; MUID:92114051

A:Accession: S26004

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-502 <ODA>

A:Cross-references: EMBL:M68929; NID:g786182; PID:g786243

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1992

C:Genetics:

A:Genome: mitochondrion

C:Keywords: mitochondrion

Query Match 66.7%; Score 32; DB 2; Length 502;

Best Local Similarity 100.0%; Pred. No. 74;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 3 VRYRY 8  
| | | | |  
Db 272 VRYRY 277

Search completed: September 7, 1999, 23:22:59  
Job time: 966 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:18 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280A-39

Perfect score: 48

Sequence: 1 RYRYRYRV 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	72.9	110	1	RK23_PORPU
2	34	70.8	675	1	KSC5_ECOLI
3	34	70.8	600	1	SPO8_YEAST
4	33	68.8	630	1	INLB_LISMO
5	33	68.8	113	1	Y109_HAEN
6	33	68.8	1095	1	YOH5_YEAST
7	33	68.8	285	1	YPDC_ECOLI
8	32	66.7	507	1	INOL_CITPA
9	32	66.7	400	1	LIPA_NEIME
10	32	66.7	868	1	MIS5_SCHPO
11	32	66.7	674	1	WHIT_HUMAN
12	32	66.7	666	1	WHIT_MOUSE
13	32	66.7	208	1	Y379_MARJA
14	32	66.7	502	1	YM40_MARPO
15	32	66.7	129	1	YRL1_SULAC
16	31	64.6	299	1	CC16_SCHPO
17	31	64.6	813	1	FGR2_XENLA
18	31	64.6	730	1	PLOD_CABEL
19	30	62.5	20	1	AMP_FUSNU
20	30	62.5	1603	1	AKO1_EMENI
21	30	62.5	511	1	CIKD_HUMAN
22	30	62.5	511	1	CIKD_MOUSE
23	30	62.5	585	1	CIKD_RAT
24	30	62.5	570	1	HEMA_NDVA
25	30	62.5	571	1	HEMA_NDVH3
26	30	62.5	571	1	HEMA_NDVI
27	30	62.5	571	1	HEMA_NDVM
28	30	62.5	436	1	RBL_EUGST
29	30	62.5	163	1	TPX1_STRPN
30	30	62.5	1118	1	YDPR_HUMAN
31	30	62.5	340	1	UBPB_ECOLI
32	29	60.4	127	1	ATP2_BACP3
33	29	60.4	249	1	CAMT_EUGGU
34	29	60.4	241	1	CAMT_PETCR
35	29	60.4	247	1	CAMT_POTPM
36	29	60.4	242	1	CAMT_VITVI
37	29	60.4	1051	1	CARB_SULSO
38	29	60.4	1333	1	CC25_CANAL
39	29	60.4	351	1	COB1_METTH
40	29	60.4	448	1	CR51_BACSH
41	29	60.4	448	1	CR53_BACSH
42	29	60.4	856	1	ENV_HV2NZ
43	29	60.4	445	1	ESAL_YEAST

ALIGNMENTS

RESULT 1									
ID	RK23_PORPU	STANDARD;	PRT;	110 AA.					
AC	P51312;								
DT	01-OCT-1996 (REL. 34, CREATED)								
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)								
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)								
DE	CHLOROPLAST 50S RIBOSOMAL PROTEIN L23.								
GN	RPL23.								
OS	PORPHYRA PURPUREA.								
OG	CHLOROPLAST								
OC	EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; BANGIALES; PORPHYRA.								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-AVONPORT;								
RA	REITH M.E., MUNDHOLLAND J.;								
RT	"Complete nucleotide sequence of the Porphyra purpurea chloroplast genome."								
RL	PLANT MOL. BIOL. REP. 13:333-335(1995).								
CC	-1- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).								
CC	EMBL; U38804; GI276778; -								
DR	PROSITE; PS00050; RIBOSOMAL_L23; 1.								
DR	PFAM; PF00276; L23; 1.								
DR	MENDEL; 10322; PORPU:tpl23;1.								
KW	RIBOSOMAL PROTEIN; CHLOROPLAST; RNA-BINDING.								
SQ	SEQUENCE 110 AA; 12782 MW; 2A87FAE3 CRC32;								
Query Match 72.9%; Score 35; DB 1; Length 110;									
Best Local Similarity 55.6%; Pred. No. 1.8;									
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;									
Qy	1 RYRYRYRV 9								
	:								
Db	2 RYLKYKRYI 10								
RESULT 2									
ID	KSC5_ECOLI	STANDARD;	PRT;	675 AA.					
AC	P42217;								
DT	01-NOV-1995 (REL. 32, CREATED)								
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)								
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)								
DE	CAPSULE POLYSACCHARIDE EXPORT PROTEIN KPSC.								
GN	KPSC.								
OS	ESCHERICHIA COLI								
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-K5;								
RX	MEDLINE; 93388530.								
RA	PAZZANI C., ROSENOW C., BOULNOIS G.J., BRONNER D., JANN K.,								
RA	ROBERTS I.S.;								
RT	"Molecular analysis of region 1 of the Escherichia coli K5 antigen gene cluster: a region encoding proteins involved in cell surface								

```

RT expression of capsular polysaccharide." ;
RL J. BACTERIOL. 175:5978-5983(1993).
RN [2]
RP SEQUENCE OF 1-8 FROM N.A.
RC STRAIN-K5;
RX MEDLINE; 95180691.
RA ROSENOW C., ROBERTS I.S., JANN K.;
RT "Isolation from recombinant Escherichia coli and characterization of
RT CMP-Kdo synthetase, involved in the expression of the capsular K5
RT Polysaccharide (K-CRS)." ;
RL FEMS MICROBIOL. LETT. 125:159-164(1995).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -!- SIMILARITY: SOME, TO THE R.MELILOTI LIPOPOLYSACCHARIDE PROCESSING
CC PROTEIN LPS2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X74567; G397408; -
DR EMBL; S76943; E193729; -
KW POLYSACCHARIDE TRANSPORT; TRANSPORT.
SQ SEQUENCE 675 AA; 75771 MW; 663AD1C3 CRC32;

Query Match 70.8%; Score 34; DB 1; Length 675;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRYV 9
DB 293 YLYRCRYI 300

RESULT 3
SPOB_YEAST
ID SPOB_YEAST STANDARD; PRT; 600 AA.
AC P41833;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DE TRANSCRIPTIONAL REGULATOR SPO8
GN SPO8 OR IME4 OR YGL192W OR GL337.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
CC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA SAKURAI M.;
RL SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-GKX5;
RA CLANCY M.J.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE; 97197971.
RA COGLIEVINA M., KLIMA R., BERTANI I., DELNERI D., ZACCARIA P.,
RA BRUSCHI C.V.;
RT "Sequencing of a 40.5 kb fragment located on the left arm of
RT chromosome VII from Saccharomyces cerevisiae." ;
RL YEAST 13:55-64(1997).
CC -!- FUNCTION: POSITIVE TRANSCRIPTIONAL REGULATOR FOR IME2.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

---

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D23721; G471168; -
DR EMBL; U30859; G945046; -
DR EMBL; X91837; E203623; -
DR EMBL; Z72714; E243795; -
DR SGD; L0000863; IME4.
KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.
SQ SEQUENCE 600 AA; 69395 MW; B4A231B CRC32;

Query Match 70.8%; Score 34; DB 1; Length 600;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRYVRYV 9
DB 272 RYHYLYQYI 280

RESULT 4
INLB_LISMO
ID INLB_LISMO STANDARD; PRT; 630 AA.
AC P25147;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DE 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE INTERNALIN B PRECURSOR.
GN INLB
OS LISTERIA MONOCYTOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
CC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91292517.
RA GAILLARD J.-L., BERCHE P., FREHEL C., GOUIN E., COSSART P.;
RT "Entry of L. monocytogenes into cells is mediated by internalin, a
RT repeat protein reminiscent of surface antigens from gram-positive
RT cocci." ;
RL CELL 65:1127-1141(1991).
CC -!- FUNCTION: MEDIATES THE ENTRY OF LISTERIA MONOCYTOGENES INTO CELLS.
CC -!- SIMILARITY: TO INTERNALIN A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M67471; G149675; -
DR PIR; C39930; C39930.
DR PFAM; PF00560; LRR; 3.
KW REPEAT; SIGNAL.
FT SIGNAL 1 30
FT CHAIN 31 630
FT DOMAIN 63 238
FT REPEAT 63 84
FT REPEAT 85 106
FT REPEAT 107 128
FT REPEAT 129 150
FT REPEAT 151 172
FT REPEAT 173 194
FT REPEAT 195 216
FT REPEAT 217 238
FT DOMAIN 399 630
FT REPEAT 399 465
FT REPEAT 559 630

```

```

SQ SEQUENCE 630 AA; 71220 MW; CA2F6E58 CRC32;

Query Match 68.8%; Score 33; DB 1; Length 630;
Best Local Similarity 71.4%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRVYRVYR 7
    |||:|
Db 401 RYVKYIR 407

RESULT 5
Y109_HAEIN STANDARD; PRT; 113 AA.
AC P43943;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN HI0109 PRECURSOR.
GN HI0109.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.-P., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32696; G1573066; -.
DR TIGR; HI0109; -.
KW HYPOTHETICAL PROTEIN; SIGNAL.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 113 HYPOTHETICAL PROTEIN HI0109.
SQ SEQUENCE 113 AA; 13379 MW; 5321F58F CRC32;

Query Match 68.8%; Score 33; DB 1; Length 113;
Best Local Similarity 62.5%; Pred. No. 4.4;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRVYRVYR 9
    :|:|:|
Db 79 FVEYRVYR 86

RESULT 6
YOH5_YEAST STANDARD; PRT; 1095 AA.
ID YOH5_YEAST
AC Q08234;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DDG-GLK INTERGENIC REGION.
GN YPDC.

DE PROBABLE ATP-DEPENDENT TRANSPORTER YOL075C.
GN YOL075C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97321807.
RA TZERMIA M., KATSIOULOU C., ALEXANDRAKI D.;
RT "Sequence analysis of a 33.2 kb segment from the left arm of yeast
RT chromosome XV reveals eight known genes and ten new open reading
RT frames including homologues of ABC transporters, inositol
RT phosphatases and human expressed sequence tags.";
RL YEAST 13:583-589(1997).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z74817; E251877; -.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
KW HYPOTHETICAL PROTEIN; ATP-BINDING; TRANSMEMBRANE; GLYCOPROTEIN;
KW TRANSPORT.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 354 374 POTENTIAL.
FT TRANSMEM 406 426 POTENTIAL.
FT TRANSMEM 840 860 POTENTIAL.
FT TRANSMEM 871 891 POTENTIAL.
FT TRANSMEM 922 942 POTENTIAL.
FT TRANSMEM 949 969 POTENTIAL.
FT TRANSMEM 978 998 POTENTIAL.
FT TRANSMEM 1009 1029 POTENTIAL.
FT TRANSMEM 1068 1088 POTENTIAL.
FT NP_BIND 528 535 ATP (POTENTIAL).
FT CARBOHYD 142 142 POTENTIAL.
FT CARBOHYD 150 150 POTENTIAL.
FT CARBOHYD 172 172 POTENTIAL.
FT CARBOHYD 329 329 POTENTIAL.
FT CARBOHYD 784 784 POTENTIAL.
FT CARBOHYD 863 863 POTENTIAL.
SQ SEQUENCE 1095 AA; 123507 MW; 24679CD6 CRC32;

Query Match 68.8%; Score 33; DB 1; Length 1095;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRVYRVYR 9
    |||:|:|
Db 351 YVRWIKYI 358

RESULT 7
YPDC_ECOLI STANDARD; PRT; 285 AA.
ID YPDC_ECOLI
AC P77396;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR IN DDG-GLK INTERGENIC REGION.
GN YPDC.

```

OS ESCHERICHIA COLI.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 CC ESCHERICHIA.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE; 97426617.  
 RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
 RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
 RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
 RA MAU B., SHAO Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL SCIENCE 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,  
 RA IKEMOTO K., INADA T., ISONO K., ITOH T., KANAI K., KASAI H.,  
 RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,  
 RA MASUDA S., MIKI T., MIZOBUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,  
 RA NASHIMOTO H., NISHIO Y., OSHIMA T., SAITO N., SAMEI G., SEKI Y.,  
 RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.;  
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XILS FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000326; G1788725; -.  
 DR EMBL; D90868; G1799793; -.  
 DR ECOGENE; EG14150; YPDC.  
 DR PROSITE; PS00041; HTH\_ARAC\_FAMILY\_1; 1.  
 DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
 DR PFAM; PF00165; HTH\_2; 1.  
 DR HYPOTHETICAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING.  
 KW DNA\_BIND 200 219 H-T-H MOTIF (BY SIMILARITY).  
 FT SEQUENCE 285 AA; 32355 MW; 61FFAA84 CRC32;  
 SQ  
 Query Match 68.8%; Score 33; DB 1; Length 285;  
 Best Local Similarity 55.6%; Pred. No. 12;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RYRVYRVYV 9  
 Db 225 REIYVRVW 233  
 I:::|::|  
 RESULT 8  
 ID INOI\_CITPA STANDARD; PRT; 507 AA.  
 AC P42802;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE MYO-INOSITOL-1-PHOSPHATE SYNTHASE (EC 5.5.1.4) (IPS).  
 OS CITRUS PARADISI (GRAPEFRUIT).  
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 CC SAPINDALES; RUTACEAE; CITRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE; 95148748.  
 RA ABU-ABIED M., HOLLAND D.;  
 RT "The gene c-inol from *Citrus paradisi* is highly homologous to turl  
 RT and inol from yeast and *Spirodela* encoding for myo-inositol phosphate

synthase.";  
 RL PLANT PHYSIOL. 106:1689-1689(1994).  
 CC -!- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE - 1L-MYO-INOSITOL  
 CC 1-PHOSPHATE.  
 CC -!- COFACTOR: NAD.  
 CC -!- PATHWAY: INOSITOL BIOSYNTHESIS.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE MYO-INOSITOL-1-PHOSPHATE SYNTHASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z32632; G602565; -.  
 DR PHOSPHOLIPID BIOSYNTHESIS; INOSITOL BIOSYNTHESIS; ISOMERASE; NAD.  
 KW SEQUENCE 507 AA; 56334 MW; 60127CC0 CRC32;  
 SQ  
 Query Match 66.7%; Score 32; DB 1; Length 507;  
 Best Local Similarity 55.6%; Pred. No. 34;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RYRVYRVYV 9  
 Db 382 RWIKYVPV 390  
 I:::|::|  
 RESULT 9  
 ID LIPA\_NEIME STANDARD; PRT; 400 AA.  
 AC Q05013;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE CAPSULE POLYSACCHARIDE MODIFICATION PROTEIN LIPA.  
 GN LIPA.  
 OS NEISSERIA MENINGITIDIS.  
 OC BACTERIA; PROTEOBACTERIA; BETA SUBDIVISION; NEISSERIACEAE; NEISSERIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ISOLATE B1940 / SEROGROUP B;  
 RX MEDLINE; 93316845.  
 RA FROSCH M., MOELLER A.;  
 RT "Phospholipid substitution of capsular polysaccharides and mechanisms  
 RT of capsule formation in *Neisseria meningitidis*."  
 RL MOL. MICROBIOL. 8:483-493(1993).  
 CC -!- FUNCTION: INVOLVED IN THE PHOSPHOLIPID MODIFICATION OF THE  
 CC CAPSULAR POLYSACCHARIDE, A STRONG REQUIREMENT FOR ITS  
 CC TRANSLOCATION TO THE CELL SURFACE.  
 CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (CYTOPLASMIC  
 CC SIDE) (PROBABLE).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z13995; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S28077; S28077.  
 DR PIR; S32879; S32879.  
 KW INNER MEMBRANE; POLYSACCHARIDE TRANSPORT; TRANSPORT.  
 SQ SEQUENCE 400 AA; 45106 MW; B99FB701 CRC32;  
 Query Match 66.7%; Score 32; DB 1; Length 400;

```

Best Local Similarity 55.68; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVRYV 9
   ||:| ||
Db 134 RYLYSRYL 142

RESULT 10
MISS_SCHPO STANDARD; PRT; 868 AA.
ID MIS5_SCHPO
AC P49731;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MISS PROTEIN.
GN MIS5.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 95170112.
RA TAKAHASHI K., YAMADA H., YANAGIDA M.;
RT "Fission yeast minichromosome loss mutants mis cause lethal
   aneuploidy and replication abnormality.";
RL MOL. BIOL. CELL 5:1145-1158(1994).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
-----
DR EMBL; D31960; G829086; -.
DR PROSITE; PS00847; MCM_1; 1.
DR PROSITE; PS50051; MCM_2; 1.
DR PFAM; PF00493; MCM; 1.
DR TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;
KW DNA REPLICATION; CELL CYCLE; ATP-BINDING.
FT DOMAIN 426 633
FT NP_BIND 476 483 ATP (POTENTIAL).
SQ SEQUENCE 868 AA; 96753 MW; EAF353B0 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 868;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVRYV 7
   ||:| ||
Db 654 RYRYRV 660

RESULT 11
WHIT_HUMAN STANDARD; PRT; 674 AA.
ID WHIT_HUMAN
AC P45844;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE WHITE PROTEIN HOMOLOG (ATP-BINDING CASSETTE TRANSPORTER 8).
GN ABC8 OR WHIT1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RN SEQUENCE FROM N.A.
RP

```

```

RC TISSUE-RETINA;
RX MEDLINE; 96256850.
RA CHEN H.M., ROSSIER C., LALIOI M.D., LYNN A., CHAKRAVARTI A.,
RA PERRIN G., ANTONARAKIS S.E.;
RT "Cloning of the cDNA for a human homologue of the Drosophila white
   gene and mapping to chromosome 21q22.3.";
RL AM. J. HUM. GENET. 59:66-75(1996).
[2]
RN SEQUENCE OF 37-674 FROM N.A.
RP TISSUE-FETAL BRAIN;
RX MEDLINE; 97186700.
RA CROOP J.M., TILLER G., FLETCHER J.A., LUX M., RAAB E., GOLDENSON D.,
RA ARCINIEGAS S., SON D., WU R.;
RT "Isolation and characterization of a mammalian homolog of the
   Drosophila white gene.";
RL GENE 185:77-85(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SEVERAL TISSUES.
CC -1- ALTERNATIVE PRODUCTS: IT IS POSSIBLE THAT TWO DIFFERENT FORMS THAT
   DIFFER IN THE PRESENCE OR ABSENCE OF A 12 RESIDUE REGION ARE
   PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
   (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
-----
DR EMBL; X91249; E218444; -.
DR EMBL; U34919; G1314277; -.
DR MIM; 603076; -.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_TRAN; 1.
KW ATP-BINDING; TRANSMEMBRANE; TRANSPORT; ALTERNATIVE SPLICING.
FT NP_BIND 114 121
FT TRANSMEM 423 441 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 503 521 POTENTIAL.
FT TRANSMEM 530 551 POTENTIAL.
FT TRANSMEM 564 582 POTENTIAL.
FT TRANSMEM 646 665 POTENTIAL.
FT VARSPIC 371 382 MISSING (IN A SHORTER FORM).
FT CONFLICT 444 444 T -> A (IN REF. 2).
FT CONFLICT 529 529 R -> A (IN REF. 2).
FT CONFLICT 664 664 L -> F (IN REF. 2).
SQ SEQUENCE 674 AA; 75169 MW; 536156C1 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 674;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YRVRYV 8
   |:| |||
Db 594 YISVRY 600

RESULT 12
WHIT_MOUSE STANDARD; PRT; 666 AA.
ID WHIT_MOUSE
AC Q64343;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE WHITE PROTEIN HOMOLOG (ATP-BINDING CASSETTE TRANSPORTER 8).
GN ABC8 OR WHIT1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE; 97186700.
RX  CROOP J.M., TILLER G.E., FLETCHER J.A., LUX M.L., RAAB E.,
RA  GOLDENSON D., SON D., ARCINIEGAS S., WU R.;
RT  "Isolation and characterization of a mammalian homolog of the
RT  Drosophila white gene.";
RL  GENE 185:77-85(1997).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-DBA/2;
RX  MEDLINE; 96359154.
RA  SAVARY S., DENIZOT F., LUCIANI M.-F., MATTEI M.-G., CHIMINI G.;
RT  "Molecular cloning of a mammalian ABC transporter homologous to
RT  Drosophila white gene.";
RL  MAMM. GENOME 7:673-676(1996).
CC  -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC  -!- TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN, THYMUS, LUNG,
CC  ADRENALS, SPLEEN AND PLACENTA. LITTLE OR NO EXPRESSION IN LIVER,
CC  KIDNEY, HEART, MUSCLE OR TESTES.
CC  -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC  (ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U34920; G1314279; -
DR  EMBL; 248745; E156982; -
DR  MGd; MG1:107704; ABC8.
DR  PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR  PFAM; PF00005; ABC_tran; 1.
DR  HSP; P30533; INRE.
KW  ATP-BINDING; TRANSMEMBRANE; TRANSPORT.
FT  NP-BIND 118 125 ATP (POTENTIAL).
FT  TRANSMEM 415 433 POTENTIAL.
FT  TRANSMEM 445 465 POTENTIAL.
FT  TRANSMEM 495 513 POTENTIAL.
FT  TRANSMEM 522 543 POTENTIAL.
FT  TRANSMEM 556 574 POTENTIAL.
FT  TRANSMEM 638 657 POTENTIAL.
SQ  SEQUENCE 666 AA; 74033 MW; AB4B792F CRC32;

Query Match 66.7%; Score 32; DB 1; Length 666;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YVRYVRY 8
DB 586 YISYVRY 592
|: ||||
|: ||||

RESULT 13
I379_METJA
ID Y379_METJA STANDARD; PRT; 208 AA.
AC Q57824;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0379.
GN MJ0379.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.

RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.B., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.".
RL SCIENCE 273:1058-1073(1996).
CC -!- SIMILARITY: IN THE C-TERMINAL, TO M.JANNASCHII MJ0723.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67490; G1591084; -
DR TIGR; MJ0379; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 208 AA; 24080 MW; 5DD82FAD CRC32;

Query Match 66.7%; Score 32; DB 1; Length 208;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRYVRYV 9
DB 2 VRYMYI 8
|: ||||
|: ||||

RESULT 14
YM40_MARPO
ID YM40_MARPO STANDARD; PRT; 502 AA.
AC P38478;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 57.7 KD PROTEIN IN 18S RRNA INTRON 1 (ORF 502).
GN YMF40.
OS MARCHANTIA POLYMORPHA (LIVERWORT).
OC MITOCHONDRION.
OC EURARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; MARCHANTIOPSIDA;
OC MARCHANTIALES; MARCHANTIACEAE; MARCHANTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92114051.
RA ODA K., YAMATO K., OHTA E., NAKAMURA Y., TAKEMURA M., NOZATO N.,
RA AKASHI K., KANEKAE T., OGURA Y., KOCHI T., OHYAMA K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome.".
RL J. MOL. BIOL. 223:1-7(1992).
CC -!- SIMILARITY: TO GROUP II INTRON MATURASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M68929; G786243; -
DR PIR; S26004; S26004.
DR PFAM; PF00078; rvt; 1.
DR MENDEL; 2114; MARPOLYMF40;1. PROTEIN.
KW MITOCHONDRION; HYPOTHETICAL

```



SQ SEQUENCE 502 AA; 57722 MW; A8621A83 CRC32;

Query Match 66.7%; Score 32; DB 1; Length 502;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VRYVRY 8

|||||

Db 272 VRYVRY 277

RESULT 15

YRL1\_SULAC

ID YRL1\_SULAC STANDARD; PRT; 129 AA.

AC P38618;

DT 01-OCT-1994 (REL. 30, CREATED)

DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 15.2 KD PROTEIN IN RL46-RLX INTERGENIC REGION (AO1).

OS SULFOLOBUS ACIDOCALDARIUS.

OC ARCHAEA; CRENARCHAEOTA; SULFOLOBALES; SULFOLOBUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DSM 639;

RX MEDLINE; 95226466.

RA MOLL R., SCHAEFER G., SCHMIDTKE S.;

RT "Nucleotide sequence of a gene cluster encoding ribosomal proteins in

the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius.";

RL BIOCHIM. BIOPHYS. ACTA 1261:315-318(1995).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X77509; G453439; -.

DR EMBL; X77509; E300459; -.

DR PIR; S41964; S41964.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 129 AA; 15196 MW; E346121E CRC32;

Query Match 66.7%; Score 32; DB 1; Length 129;

Best Local Similarity 55.6%; Pred. No. 7.9;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYRVRYV 9

|||||

Db 68 RAIKRYRYM 76

Search completed: September 7, 1999, 23:59:19

Job time: 531 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:49 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: US-09-124-280A-39

Perfect score: 48

Sequence: 1 RYRVRYV 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	72.9	426	11	O55082
2	34	70.8	261	1	O26968
3	34	70.8	394	2	O85335
4	34	70.8	438	4	O75783
5	34	70.8	149	13	O9YHK6
6	33	68.8	415	2	O92739
7	33	68.8	533	4	O15317
8	33	68.8	337	5	O17433
9	33	68.8	1928	5	O97291
10	33	68.8	325	12	O11441
11	32	66.7	322	2	O31656
12	32	66.7	481	2	P71915
13	32	66.7	306	3	O74870
14	32	66.7	900	3	O74925
15	32	66.7	232	4	O43576
16	32	66.7	200	5	O20126
17	32	66.7	425	5	O18923
18	32	66.7	296	5	O93456
19	32	66.7	165	5	O96406
20	32	66.7	785	8	O33759
21	32	66.7	636	8	O39479
22	32	66.7	415	10	O81088
23	32	66.7	857	12	O66956
24	32	66.7	852	12	O66959
25	31	64.6	223	1	O58882
26	31	64.6	645	2	O45773
27	31	64.6	170	2	P76492
28	31	64.6	301	2	O07156
29	31	64.6	514	2	O67596

30	31	64.6	346	2	O69652
31	31	64.6	1166	2	O92757
32	31	64.6	829	5	O18253
33	31	64.6	414	5	O45571
34	31	64.6	1259	8	O35058
35	31	64.6	336	10	O39218
36	31	64.6	391	10	O42532
37	31	64.6	396	12	O56869
38	30	62.5	228	2	O34903
39	30	62.5	249	2	O83299
40	30	62.5	280	2	O85861
41	30	62.5	37	2	O49100
42	30	62.5	238	2	O45591
43	30	62.5	862	5	O22354
44	30	62.5	477	5	O19988
45	30	62.5	1033	5	O44323

ALIGNMENTS

RESULT 1

O55082  
ID O55082 PRELIMINARY; PRT; 426 AA.  
AC O55082;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 20 (EC 3.1.3.48)  
DE (PHOSPHOTYROSINE PHOSPHATASE) (PTPASE).  
GN PTPN20.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-TESTIS;  
RX MEDLINE; 98070510.  
RA OHSUGI M., KURAMUCHI S., MATSUDA S., YAMAMOTO T.;  
RT "Molecular cloning and characterization of a novel cytoplasmic  
protein-tyrosine phosphatase that is specifically expressed in  
spermatocytes.";  
RL J. Biol. Chem. 272:33092-33099(1997).  
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN  
CC TYROSINE + PHOSPHATE.  
DR EMBL; D64141; BAA23761.1; -;  
DR MGD; MG1:1196295; PTPN20.  
DR PFAM; PF00102; Y\_Posphatase; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
KW Hydrolase.  
SQ SEQUENCE 426 AA; 49118 MW; 58F1AA05 CRC32;

Query Match 72.9%; Score 35; DB 11; Length 426;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 YRVRYV 9  
Db 339 FIKVRYV 346

RESULT 2

O26968  
ID O26968 PRELIMINARY; PRT; 261 AA.  
AC O26968;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
DE CONSERVED PROTEIN.  
GN MTH882.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;

```

OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTRIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000864; AAB85380.1; -.
SQ SEQUENCE 261 AA; 26929 MW; 53F0B865 CRC32;

Query Match 70.8%; Score 34; DB 1; Length 261;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRYRVY 9
Db 241 VYRIRY 247
| | | | |
| | | | |

RESULT 3
O85335 PRELIMINARY; PRT; 394 AA.
AC O85335;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE O ANTIGEN POLYMERASE WZY.
GN WZY.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C664-1992;
RX MEDLINE; 98339851.
RA WANG L., REEVES P.R.;
RT "Organization of Escherichia coli O157 O antigen gene cluster and
RT identification of its specific genes.";
RL Infect. Immun. 66:3545-3551(1998).
DR EMBL; AF061251; AAC32340.1; -.
SQ SEQUENCE 394 AA; 45647 MW; EF0C67FB CRC32;

Query Match 70.8%; Score 34; DB 2; Length 394;
Best Local Similarity 57.1%; Pred. No. 65;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRVYRV 8
Db 117 YRIRY 123
| | | | |
| | | | |

RESULT 4
O75783 PRELIMINARY; PRT; 438 AA.
ID O75783;
AC O75783;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE RHOMBOID-RELATED PROTEIN.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE; 98324821.
RA PASCALL J.C., BROWN K.D.;
RT "Characterization of a mammalian cDNA encoding a protein with high
RT sequence similarity to the Drosophila regulatory protein Rhomboid.";
RL FEBS Lett. 429:337-340(1998).
DR EMBL; Y17108; CAA76629.1; -.
SQ SEQUENCE 438 AA; 48314 MW; E05F94D5 CRC32;

Query Match 70.8%; Score 34; DB 4; Length 438;
Best Local Similarity 75.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYRVYRVY 8
Db 168 RYRVYRVY 175
| | | | |
| | | | |

RESULT 5
O9YHK6 PRELIMINARY; PRT; 149 AA.
ID O9YHK6;
AC O9YHK6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE ANTIGEN RECEPTOR (FRAGMENT).
GN NAR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Orectolobiformes; Ginglymostomatidae; Ginglymostoma.
RN [1]
RP SEQUENCE FROM N.A.
RA DIAZ M., GRENNBERG A.S., FLAJNIK M.F.;
RT "Somatic Hypermutation of the New Antigen Receptor Gene (NAR) in the
RT Nurse Shark Does Not Generate the Repertoire: Possible Role in
RT Antigen-Driven Reactions in the Absence of Germinal Centers.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14343-14348(1998).
DR EMBL; AF096029; AAC83731.1; -.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 149 149
SQ SEQUENCE 149 AA; 16362 MW; 8F1AD358 CRC32;

Query Match 70.8%; Score 34; DB 13; Length 149;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYRVYRVY 9
Db 53 RYVETRVY 61
| | | | |
| | | | |

RESULT 6
O9Z739 PRELIMINARY; PRT; 415 AA.
ID O9Z739;
AC O9Z739;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE ROD SHAPE PROTEIN.
GN RODA.
OS Chlamydia pneumoniae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RA KALMAN S., MITCHELL W., MARATHE R., LAMMEL C., FAN J., OLINGER L.,
RA GRIMWOOD J., DAVIS R.W., STEPHENS R.S.;
RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001667; AAD19005.1; -.

```

SQ SEQUENCE 415 AA; 46662 MW; 3A4ABBC4 CRC32;

Query Match 68.8%; Score 33; DB 2; Length 415;

Best Local Similarity 66.7%; Pred. No. 1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RYRYRYRYV 9

Db 38 RYRYRYRYV 46

RESULT 7

ID O15317 PRELIMINARY; PRT; 533 AA.

AC O15317;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE RNA POLYMERASE III SUBUNIT.

GN RPC62.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 97315201.

RA WANG Z.; ROEDER R.G.;

RT "Three human RNA polymerase III-specific subunits form a subcomplex

with a selective function in specific transcription initiation.";

RL Genes Dev. 11:1315-1326(1997).

DR EMBL; U93867; AAB63675.1; -.

SQ SEQUENCE 533 AA; 60629 MW; 851C71CF CRC32;

Query Match 68.8%; Score 33; DB 4; Length 533;

Best Local Similarity 66.7%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RYRYRYRYV 9

Db 84 RYRYRYRYV 92

RESULT 8

ID Q17433 PRELIMINARY; PRT; 337 AA.

AC Q17433;

DT 01-JAN-1999 (TrEMBLrel. 09, Created)

DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)

DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)

DE HYPOTHETICAL 40.0 KD PROTEIN B0035.2 IN CHROMOSOME IV PRECURSOR.

GN B0035.2

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;

OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA WHITE S.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS A DNAJ-LIKE DOMAIN.

DR EMBL; Z73102; CAA97409.1; -.

DR WORMPEP; B0035.2; CF05161.

DR PROSITE; PS00636; DNAJ\_1; 1.

DR PFAM; PF00226; DnaJ; 1.

DR Hypothetical protein; Signal; Chaperone.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 337 HYPOTHETICAL PROTEIN B0035.2.

FT DOMAIN 34 107 DNAJ-LIKE.

SQ SEQUENCE 337 AA; 39992 MW; 5A8733EC CRC32;

Query Match 68.8%; Score 33; DB 5; Length 337;

Best Local Similarity 66.7%; Pred. No. 83;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RYRYRYRYV 9

Db 317 RYRYRYRYV 325

RESULT 9

ID O97291 PRELIMINARY; PRT; 1928 AA.

AC O97291;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

DE MAL3P7.21 PROTEIN.

GN MAL3P7.21.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-3D7;

RA OLIVER K.; BOWMAN S.; HARRIS D.; LAWSON D.; QUAIL M.; BARRELL B.;

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL034559; CAB39037.1; -.

SQ SEQUENCE 1928 AA; 230190 MW; A8AE8DEB CRC32;

Query Match 68.8%; Score 33; DB 5; Length 1928;

Best Local Similarity 75.0%; Pred. No. 4.7e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 YRYRYRYV 9

Db 964 YRYRYRYV 971

RESULT 10

ID O11441 PRELIMINARY; PRT; 325 AA.

AC O11441;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)

DE RNA REPLICASE (FRAGMENT).

GN PCL.

OS Urochloa hoja blanca virus.

OC Viruses; ssRNA negative-strand viruses; Tenuivirus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-COSTA RICA;

RA DE MIRANDA J.R.; MUNOZ M.; CABEZAS E.; HERNANDEZ M.; ROOSINCK M.J.;

RA WU R.; ESPINOZA A.M.;

RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; U82448; AAB58306.1; -.

FT NON\_TER 1

SQ SEQUENCE 325 AA; 37992 MW; 80D2678A CRC32;

Query Match 68.8%; Score 33; DB 12; Length 325;

Best Local Similarity 85.7%; Pred. No. 80;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 3 YRYRYRYV 9

Db 260 YRYRYRYV 266

RESULT 11

O51656

ID O51656 PRELIMINARY; PRT; 322 AA.

AC O51656;

DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 38.6 KD PROTEIN.  
 GN BB0714.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE; 98065943.  
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA OTTERBACK T., WATHEY L., McDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi".  
 RL Nature 390:580-586(1997).  
 DR EMBL; AF001171; AAC67064.1; -.  
 DR TIGR; BB0714; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 322 AA; 38625 MW; 79D82CBA CRC32;  
  
 Query Match 66.7%; Score 32; DB 2; Length 322;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 RYRVYRVY 8  
 DB 4 RYLRVIFY 11  
  
 RESULT 12  
 ID P71915 PRELIMINARY; PRT; 481 AA.  
 AC P71915;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 52.3 KD PROTEIN.  
 GN MTCY428.12.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA DEVLIN K., CHURCHER C.M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., HILBERT H., DUESTERHOEFF A;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031855; CAA21292.1; -.  
 SQ SEQUENCE 900 AA; 103719 MW; F96A69E7 CRC32;  
  
 Query Match 66.7%; Score 32; DB 3; Length 306;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 1 RYRVYRVY 9  
 DB 96 RMVYVRYI 104  
  
 RESULT 14  
 ID O74925 PRELIMINARY; PRT; 900 AA.  
 AC O74925;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE PUTATIVE VACUOLAR MEMBRANE PROTEIN.  
 GN SPC3790.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., HILBERT H., DUESTERHOEFF A;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031855; CAA21292.1; -.  
 SQ SEQUENCE 900 AA; 103719 MW; F96A69E7 CRC32;  
  
 Query Match 66.7%; Score 32; DB 3; Length 900;  
 Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 3 RYRVYRVY 9  
 DB 629 IRYLRVY 635

SQ SEQUENCE 481 AA; 52272 MW; F54759A9 CRC32;  
  
 Query Match 66.7%; Score 32; DB 2; Length 481;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 1 RYRVYRVY 8  
 DB 348 RYARLVRY 355  
  
 RESULT 13  
 ID O74870 PRELIMINARY; PRT; 306 AA.  
 AC O74870;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN SPC31H12.02C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA WOOD V., RAJANDREAM M.A., BARRELL B.G., RIEGER M;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031824; CAA21219.1; -.  
 DR PFAM; PF01036; Bac\_rhodopsin; 1.  
 SQ SEQUENCE 306 AA; 34663 MW; 16578FD7 CRC32;  
  
 Query Match 66.7%; Score 32; DB 3; Length 306;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 1 RYRVYRVY 9  
 DB 96 RMVYVRYI 104  
  
 RESULT 14  
 ID O74925 PRELIMINARY; PRT; 900 AA.  
 AC O74925;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE PUTATIVE VACUOLAR MEMBRANE PROTEIN.  
 GN SPC3790.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., HILBERT H., DUESTERHOEFF A;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031855; CAA21292.1; -.  
 SQ SEQUENCE 900 AA; 103719 MW; F96A69E7 CRC32;  
  
 Query Match 66.7%; Score 32; DB 3; Length 900;  
 Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 3 RYRVYRVY 9  
 DB 629 IRYLRVY 635

```

RESULT 15
O43576 PRELIMINARY; PRT; 232 AA.
ID O43576;
AC O43576;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE WHITE PROTEIN HOMOLOG (FRAGMENT).
GN WHITE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE: 96207227.
RA ANDERSSON B., WENTLAND M.A., RICAFRENTE J.Y., LIU W., GIBBS R.A.;
RT "A "double adaptor" method for improved shotgun library
RT construction.";
RL Anal. Biochem. 236:107-113(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE: 97264341.
RA YU W., ANDERSSON B., WORLEY K.C., MUZNY D.M., DING Y., LIU W.,
RA RICAFRENTE J.Y., WENTLAND M.A., LENNON G., GIBBS R.A.;
RT "Large-scale concatenation cDNA sequencing.";
RL Genome Res. 7:353-358(1997).
DR EMBL; AF038175; AAB97364.1; -.
FT NON_TER
SQ SEQUENCE 232 AA; 26479 MW; 9E06238B CRC32;

```

```

Query Match 66.78; Score 32; DB 4; Length 232;
Best Local Similarity 71.48; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YRVYVRY 8
Db 152 YISVRY 158

```

Search completed: September 7, 1999, 22:47:50  
Job time: 7970 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:09 ; Search time 135.78 Seconds  
(without alignments)  
1.570 Million cell updates/sec

Title: US-09-124-280A-40

Perfect score: 51  
Sequence: 1 KFFKFFKFF 9

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	10	1 R71781	Peptide neutralisi
2	51	100.0	10	1 W21598	Antibiotic potenti
3	45	88.2	9	1 W21628	Antibiotic potenti
4	39	76.5	15	1 R57402	Peptide for treati
5	36	70.6	18	1 P90327	Alpha-helix-formin
6	34	66.7	99	1 P50339	IFN-gamma-(48-148)
7	34	66.7	17	1 R57403	Peptide for treati
8	34	66.7	15	1 R57404	Peptide for treati
9	34	66.7	12	1 R57387	Peptide for treati
10	34	66.7	323	1 W15759	Cotton fibrous tis
11	34	66.7	8	1 W71166	Peptide used in an
12	33	64.7	32	1 R33514	Generic peptide fo
13	33	64.7	11	1 R33532	Peptide for treati
14	33	64.7	11	1 R39294	Endotoxin lipid A
15	33	64.7	10	1 R71779	Peptide neutralisi
16	33	64.7	12	1 R71791	Peptide neutralisi
17	33	64.7	236	1 W21024	H. pylori cell env
18	33	64.7	11	1 W21623	Antibiotic potenti
19	33	64.7	10	1 W21629	Antibiotic potenti
20	33	64.7	10	1 W21596	Antibiotic potenti
21	33	64.7	12	1 W21608	Antibiotic potenti
22	33	64.7	384	1 W55338	H. pylori ORF hp4p
23	33	64.7	376	1 W55627	H. pylori ORF hp6p
24	33	64.7	166	1 W83904	Porcine interferon
25	33	64.7	166	1 W73233	Porcine interferon
26	33	64.7	382	1 W98312	H. pylori GHPO 886
27	33	64.7	222	1 Y10969	H. pylori ORF hp5p
28	32	62.7	3135	1 R57474	P. falciparum tran
29	31	60.8	132	1 P90362	Recombinant Group
30	31	60.8	887	1 R11988	Dictyostelium prot
31	31	60.8	463	1 W04360	Human cathepsin C
32	31	60.8	566	1 W20945	H. pylori inner me
33	31	60.8	13	1 W13810	Antimicrobial cati
34	31	60.8	13	1 W12888	Antimicrobial cati
35	31	60.8	439	1 W34456	Bovine dipeptidyla
36	31	60.8	463	1 W34457	Bovine dipeptidyla
37	31	60.8	374	1 W63694	Human secreted pro
38	31	60.8	1969	1 W72419	Rice bacterial lea
39	31	60.8	63	1 W98496	H. pylori GHPO 109
40	31	60.8	445	1 W98337	H. pylori GHPO 142
41	31	60.8	574	1 Y00940	Aquifex VF-5 DNA p
42	30	58.8	22	1 R10643	Magainin II analog
43	30	58.8	566	1 R13228	Endoglucanase enco

ALIGNMENTS

RESULT 1

ID R71781  
AC R71781 standard; peptide; 10 AA.  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 12; Page 21; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)<sub>n</sub>, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (A)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)<sub>p</sub>, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.003;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9

|||||||

Db 1 KFFKFFKFF 9

RESULT 2

W21598

ID W21598 standard; peptide; 10 AA.

AC W21598;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #10.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

PN W09638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic

PT amino acid sequence - reduces dose of antibiotic required

PS Claim 14; Page 25; 37pp; English.

CC The sequences given in W21598-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 51; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.003; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

QY 1 KFFKFFKFF 9  
 |||||  
 Db 1 KFFKFFKFF 9

RESULT 3  
 W21628  
 ID W21628 standard; peptide; 9 AA.  
 AC W21628;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #40.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-Al.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03;  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 44; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 9 AA;

Query Match 88.2%; Score 45; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-05; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0;

QY 1 KFFKFFKFF 8  
 |||||  
 Db 1 KFFKFFKFF 8

RESULT 4  
 R57402  
 ID R57402 standard; Protein; 15 AA.  
 AC R57402;  
 DT 21-MAR-1995 (first entry)  
 DE Peptide for treating diseases related to anti-DNA antibodies.  
 KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
 OS Synthetic.  
 PN J06192290-A.  
 PD 12-JUL-1994.  
 PF 18-JAN-1993; 006098.  
 PR 30-SEP-1992; JP-261821.  
 PA (KURS) KURARAY CO LTD.  
 DR WPI; 94-260510/32.  
 PT A peptide and an adsorbing agent prep'd. by immobilising it on a  
 PT carrier - useful for treatment of diseases related to anti-DNA

PT antibodies and immune complexes  
 PS Disclosure; Page 12; 14pp; Japanese.  
 CC The sequences given in R57386-413 are peptides which are all covered  
 CC by the claimed generic formula:  
 CC H-X-(A-B)n-Y-Z  
 CC A = trp, Phe or a peptide fragment consisting of 2 residues;  
 CC B = trp, Phe, Asn or Glu;  
 CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
 CC consisting of 2-10 residues, provided that at least one of  
 CC X or Y are present;  
 CC Z = OH or NH<sub>2</sub>; and  
 CC n = 2-5.

CC These peptides may be immobilised on a carrier in the preparation of an  
 CC absorbing agent which may be used in the treatment of diseases related  
 CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 15 AA;

Query Match 76.5%; Score 39; DB 1; Length 15;  
 Best Local Similarity 77.8%; Pred. No. 0.38; Mismatches 2; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

QY 1 KFFKFFKFF 9  
 |||||  
 Db 2 KFFDFEDEF 10

RESULT 5  
 P90327  
 ID P90327 standard; peptide; 18 AA  
 AC P90327  
 DT 16-MAR-1990 (first entry)  
 DE Alpha-helix-forming peptide emulsifier with hydrophobic and hydrophilic  
 KW Polypeptide emulsifier; hydrophobic axial domain; hydrophilic  
 KW axial domain; alpha-helix.  
 PN W08904209-A.  
 PD 18-MAY-1989.

PF 3-NOV-1988; G00969.  
 PR 01-JUL-1988; GB-015748; GB-026132.  
 PA (AGRI-) Agric & Food Res.  
 PI Brock CJ, Enser MB;  
 DR WPI; 89-165540/22.  
 PT Polypeptide(s) having alpha-helix forming amino acid sequences  
 PT - having hydrophilic and hydrophobic domains, used as emulsifiers esp.  
 PT in foods.  
 PS Claim 11 (28); 47pp; English.  
 CC domains.

CC It is a peptide emulsifier which is based on Peptide I (p90300) in which  
 CC all unchanged side chains are aromatic. It is one of a series of peptide  
 CC emulsifiers which are harmless when taken orally and have some nutritional  
 CC value and which can be used as an emulsifier in the mfr. of foods,  
 CC cosmetics, pharmaceutical preparations, cleansing formulations,  
 CC biosensors and thermodynamically-stable microemulsions.  
 SQ Sequence 18 AA;

Query Match 70.6%; Score 36; DB 1; Length 18;  
 Best Local Similarity 62.5%; Pred. No. 1.4; Mismatches 3; Indels 0; Gaps 0;  
 Matches 5; Conservative 3;

QY 2 FFKFFKFF 9  
 I:||||  
 Db 6 FYKFFEYF 13

RESULT 6  
 P50399  
 ID P50399 standard; Protein; 99 AA.  
 AC P50399;  
 DT 16-MAR-1992 (first entry)  
 DE IFN-gamma-(48-148).  
 KW Interferon; homo sapiens; human.  
 OS Synthetic.

PN J60041697-A.  
PD 05-MAR-1985.  
PF 15-AUG-1983; 148026.  
PR 15-AUG-1983; JP-148026.  
PA (ASAH ) ASahi CHEMICAL IND KK.  
DR WPI: 85-090995/15.  
DR N-PSDB; N50456.  
PT Synthesis of active protein derivs. - by reacting two peptide(s)  
PT one of which is synthetically formed by DNA recombination.  
PS Disclosure; Page 17; 23pp; Japanese.  
CC The sequence was prepd. from 25 synthetic oligonucleotides which  
CC were annealed to a further 25 making up the complementary strand,  
CC to form a ds duplex. The sequence comprises residues 48-146 of  
CC IFN gamma. An Ile residue (no. 47) was later added to the N-  
CC terminal of the expressed peptide. A second duplex encoding  
CC residues 1-46 of the protein was also synthesised (see N50455).  
CC The two peptides were expressed, purified and then reacted in the  
CC presence of a condensing agent to prepare the complete peptide.  
CC Alternatively one of the peptides is converted to its reactive  
CC deriv. and then reacted with the other peptide.  
CC See also P50398 and P50400.  
CC Sequence 99 AA;  
SQ

Query Match 66.7%; Score 34; DB 1; Length 99;  
Best Local Similarity 75.0%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFKFKF 9  
DB 9 YFKFKNF 16  
||| |||

RESULT 7  
R57403  
ID R57403 standard; Protein; 17 AA.  
AC R57403;  
DT 21-MAR-1995 (first entry)  
DE Peptide for treating diseases related to anti-DNA antibodies.  
KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
OS Synthetic.  
PN J06192290-A.  
PD 12-JUL-1994.  
PF 18-JAN-1993; 006098.  
PR 30-SEP-1992; JP-261821.  
PA (KURS ) KURARAY CO LTD.  
DR WPI: 94-260510/32.  
PT A peptide and an adsorbing agent prepd. by immobilising it on a  
PT carrier - useful for treatment of diseases related to anti-DNA  
PT antibodies and immune complexes  
PS Disclosure; Page 12; 14pp; Japanese.  
CC The sequences given in R57386-413 are peptides which are all covered  
CC by the claimed generic formula:  
CC H-X-(A-B)n-Y-Z  
CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
CC B = Trp, Phe, Asn or Glu;  
CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
CC consisting of 2-10 residues, provided that at least one of  
CC X or Y are present;  
CC Z = OH or NH2; and  
CC n = 2-5.  
CC These peptides may be immobilised on a carrier in the preparation of an  
CC adsorbing agent which may be used in the treatment of diseases related  
CC to anti-DNA antibodies and/or immune complex.  
CC Sequence 17 AA;  
SQ

Query Match 66.7%; Score 34; DB 1; Length 17;  
Best Local Similarity 75.0%; Pred. No. 2.8;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFKF 9  
DB 9 YFKFKNF 16  
||| |||

Db 6 FFDFDFDF 13

RESULT 8  
R57404  
ID R57404 standard; Protein; 15 AA.  
AC R57404;  
DT 21-MAR-1995 (first entry)  
DE Peptide for treating diseases related to anti-DNA antibodies.  
KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
OS Synthetic.  
PN J06192290-A.  
PD 12-JUL-1994.  
PF 18-JAN-1993; 006098.  
PR 30-SEP-1992; JP-261821.  
PA (KURS ) KURARAY CO LTD.  
DR WPI: 94-260510/32.  
PT A peptide and an adsorbing agent prepd. by immobilising it on a  
PT carrier - useful for treatment of diseases related to anti-DNA  
PT antibodies and immune complexes  
PS Disclosure; Page 13; 14pp; Japanese.  
CC The sequences given in R57386-413 are peptides which are all covered  
CC by the claimed generic formula:  
CC H-X-(A-B)n-Y-Z  
CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
CC B = Trp, Phe, Asn or Glu;  
CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
CC consisting of 2-10 residues, provided that at least one of  
CC X or Y are present;  
CC Z = OH or NH2; and  
CC n = 2-5.  
CC These peptides may be immobilised on a carrier in the preparation of an  
CC adsorbing agent which may be used in the treatment of diseases related  
CC to anti-DNA antibodies and/or immune complex.  
CC Sequence 15 AA;  
SQ

Query Match 66.7%; Score 34; DB 1; Length 15;  
Best Local Similarity 75.0%; Pred. No. 2.5;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFKF 9  
DB 2 FFDFDFDF 9  
||| |||

RESULT 9  
R57387  
ID R57387 standard; Protein; 12 AA.  
AC R57387;  
DT 21-MAR-1995 (first entry)  
DE Peptide for treating diseases related to anti-DNA antibodies.  
KW Carrier; absorbing agent; treatment; anti-DNA antibody; immune complex.  
OS Synthetic.  
PN J06192290-A.  
PD 12-JUL-1994.  
PF 18-JAN-1993; 006098.  
PR 30-SEP-1992; JP-261821.  
PA (KURS ) KURARAY CO LTD.  
DR WPI: 94-260510/32.  
PT A peptide and an adsorbing agent prepd. by immobilising it on a  
PT carrier - useful for treatment of diseases related to anti-DNA  
PT antibodies and immune complexes  
PS Disclosure; Page 11; 14pp; Japanese.  
CC The sequences given in R57386-413 are peptides which are all covered  
CC by the claimed generic formula:  
CC H-X-(A-B)n-Y-Z  
CC A = Trp, Phe or a peptide fragment consisting of 2 residues;  
CC B = Trp, Phe, Asn or Glu;  
CC X and Y = a bond or Asp, Glu, Arg, Lys, His or a peptide fragment  
CC consisting of 2-10 residues, provided that at least one of  
CC X or Y are present;  
CC Z = OH or NH2; and  
CC n = 2-5.  
CC These peptides may be immobilised on a carrier in the preparation of an  
CC adsorbing agent which may be used in the treatment of diseases related  
CC to anti-DNA antibodies and/or immune complex.  
CC Sequence 12 AA;  
SQ

Query Match 66.7%; Score 34; DB 1; Length 12;  
Best Local Similarity 75.0%; Pred. No. 2.5;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFKF 9  
DB 2 FFDFDFDF 9  
||| |||

CC n = 2-5.  
 CC These peptides may be immobilised on a carrier in the preparation of an  
 CC absorbing agent which may be used in the treatment of diseases related  
 CC to anti-DNA antibodies and/or immune complex.  
 SQ Sequence 12 AA;

Query Match 66.7%; Score 34; DB 1; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 2.1;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFKFKFF 9  
 || || ||  
 Db 1 FFDFFDF 8

## RESULT 10

W15759  
 ID W15759 standard; Protein; 323 AA.  
 AC W15759;  
 DT 19-JUN-1997 (first entry)  
 DE Cotton fibrous tissue specific protein KC18.  
 KW Fibrous tissue; cotton fibre; cotton flower; Gossypium barbadense;  
 KW Gossypium hirsutum.  
 OS Gossypium barbadense.  
 PN J09075093-A.  
 PD 25-MAR-1997.  
 PF 20-FEB-1996; 031987.  
 PR 21-FEB-1995; US-391696.  
 PR 29-DEC-1995; US-580545.  
 PA (TOYM ) TOYOBO KK.  
 PA (UYTE-) UNIV TEXAS TECH.  
 DR WPI: 97-239274/22.  
 DR N-PSDB: T60567.  
 PT Cotton fibrous tissue gene - used to produce transformants with  
 PT improved cotton fibres, and improved yield  
 PS Claim 3; Page 18-19; 37pp; Japanese.  
 CC W15759-W15763 represent cotton fibrous tissue proteins of the invention.  
 CC The genes encoding these sequences are specifically expressed during the  
 CC growth of cotton fibre. The genes were isolated from the polyA RNA of the  
 CC ovules of a cotton flower of Gossypium barbadense, and G. hirsutum. The  
 CC genes encoding these sequences, and vectors containing them are used in  
 CC the preparation of cotton transformants, with improved cotton fibres and  
 CC improved yield.  
 SQ Sequence 323 AA;

Query Match 66.7%; Score 34; DB 1; Length 323;  
 Best Local Similarity 66.7%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KFKFKFF 9  
 || || || || ||  
 Db 251 KLLKFFSFF 259

## RESULT 11

W11166  
 ID W11166 standard; peptide; 8 AA.  
 AC W11166;  
 DT 26-OCT-1998 (first entry)  
 DE Peptide used in an absorption material for hyperlipidemia therapy.  
 KW Absorption material; hyperlipidemia therapy; removal; LDL; blood.  
 OS Synthetic.  
 PN J10179733-A.  
 PD 07-JUL-1998.  
 PF 20-DEC-1996; 354650.  
 PR 20-DEC-1996; JP-354650.  
 PA (ASAH ) ASAH KASEI KOGYO KK.  
 PA (ASAH ) ASAH MEDICAL CO LTD.  
 DR WPI: 98-440253/38.  
 PT Absorption material for hyperlipidaemia therapy - comprises peptide  
 PT containing amino acid satisfying predetermined formulae

PS Example 2; Page 9; 11pp; Japanese.  
 CC W11164-67 represent peptides of electric charge used in an absorption  
 CC material for hyperlipidemia therapy. The material enables the safe  
 CC removal of LDL from blood.  
 SQ Sequence 8 AA;

Query Match 66.7%; Score 34; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFKFFK 7  
 || || || ||  
 Db 1 FFKFFK 6

## RESULT 12

R33514  
 ID R33514 standard; peptide; 32 AA.  
 AC R33514;  
 DT 07-JUL-1993 (first entry)  
 DE Generic peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

Query Match 64.7%; Score 33; DB 1; Length 32;  
 Best Local Similarity 75.0%; Pred. No. 7.1;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFKFFKFF 8  
 || || || ||  
 Db 2 KFLKFLKF 9

## RESULT 13

R33532  
 ID R33532 standard; peptide; 11 AA.  
 AC R33532;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 OS Synthetic.

PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 10; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of formula R-(Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val)<sub>n</sub>-R, where n = 1-100 and each R is H, an amino acid residue or a fatty acid residue. The peptide is useful for treating or preventing septic shock, mixing with polymyxin B to reduce its toxicity; removing endotoxins from blood, sera or other fluids (in vivo or in vitro); controlling release of cytokines induced by endotoxins; as diagnostic reagents to detect and quantify toxins in blood or sera; preparing non-toxic antigenic complexes of lipid A or lipopolysaccharide (LPS); and for treating pertussis, bacterial meningitis and HIV-related infections. The usual dose is 10-100 ug/kg/day, given parenterally. It binds to the same sites as polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It has no antibiotic activity; does not lyse erythrocytes; has no toxicity in mice when injected at 50mg/kg and is relatively unstable against proteases.  
 CC Sequence 11 AA;  
 SQ

Query Match 64.7%; Score 33; DB 1; Length 11;  
 Best Local Similarity 75.0%; Pred. NO. 2.8;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KFFKFFKF 8  
 || || || ||  
 Db 2 KFLKFLKF 9

RESULT 14  
 R39294

ID R39294 standard; peptide; 11 AA.  
 AC R39294;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 KW quantification.  
 OS Synthetic.  
 PN WO9314115-A.  
 PD 22-JUL-1993.  
 PF 14-MAY-1992; E01060.  
 PR 16-JAN-1992; US-819893.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin - forming non-toxic, antigenic complex, used to treat or prevent septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A component of bacterial endotoxin at the same site as polymyxin B (PMB) and with about the same affinity to produce a non-toxic antigenic complex. Unlike PMB it is not toxic, is susceptible to proteolytic degradation in serum, has no antibiotic activity and no haemolytic action. It is especially used to treat or prevent septic shock; to reduce toxicity of PMB; to remove endotoxins from blood, sera, vaccines, drug solns. etc.; to control release of cytokines induced by endotoxins; for in vivo or in vitro detoxification of bacterial endotoxins, and to detect or quantify endotoxins in blood products.

SQ Sequence 11 AA;

Query Match 64.7%; Score 33; DB 1; Length 11;  
 Best Local Similarity 75.0%; Pred. NO. 2.8;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KFFKFFKF 8  
 || || || ||  
 Db 2 KFLKFLKF 9

RESULT 15  
 R71779

ID R71779 standard; peptide; 10 AA.  
 AC R71779;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN WO9503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI; 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids  
 PT Claim 10; Page 20; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides of formula: (A)<sub>n</sub>, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)<sub>p</sub>, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.  
 CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new peptides.  
 CC Sequence 10 AA;

Query Match 64.7%; Score 33; DB 1; Length 10;  
 Best Local Similarity 75.0%; Pred. NO. 2.6;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KFFKFFKF 8  
 || || || ||  
 Db 1 KFLKFLKF 8

Search completed: September 7, 1999, 22:50:10  
 Job time: 7891 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:07:51 ; Search time 82.37 seconds  
(without alignments)  
1.078 Million cell updates/sec

Title: US-09-124-280A-40

Perfect score: 51

Sequence: 1 KFFKFFKFF 9

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/PTUS9\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	10	1	US-08-097-830E-10
2	51	100.0	10	2	US-08-456-112B-10
3	51	100.0	9	2	US-08-456-112B-40
4	33	64.7	11	1	US-08-049-871-8
5	33	64.7	11	1	US-07-819-893-8
6	33	64.7	11	1	US-08-280-397-8
7	33	64.7	11	1	US-08-097-830E-8
8	33	64.7	12	1	US-08-097-830E-20
9	33	64.7	10	2	US-08-456-112B-8
10	33	64.7	12	2	US-08-456-112B-20
11	33	64.7	11	2	US-08-456-112B-37
12	33	64.7	10	2	US-08-456-112B-41
13	32	62.7	3135	1	US-08-323-170B-2
14	31	60.8	887	1	US-07-867-106-3
15	31	60.8	463	1	US-08-426-428-2
16	31	60.8	350	1	US-08-415-751-16
17	30	58.8	8	1	US-08-257-782-25
18	30	58.8	920	1	US-08-101-593-2
19	30	58.8	920	1	US-08-101-593-4
20	30	58.8	8	2	US-08-577-846-25
21	30	58.8	78	2	US-08-680-726A-60
22	29	56.9	775	1	US-07-603-133B-14
23	29	56.9	525	1	US-08-077-939-19
24	29	56.9	525	1	US-08-461-599-19
25	29	56.9	525	1	US-08-461-621-19
26	29	56.9	525	1	US-08-462-334-19
27	29	56.9	208	2	US-08-462-169B-20
28	29	56.9	1434	2	US-08-540-406-10
29	29	56.9	1447	2	US-08-540-406-19
30	29	56.9	1434	3	PCT-US95-13233-10
31	29	56.9	1447	3	PCT-US95-13233-19
32	28	54.9	146	1	US-08-030-077-1
33	28	54.9	122	1	US-08-354-456A-9
34	28	54.9	143	2	US-08-449-639-8
35	28	54.9	418	3	PCT-US96-10986-18
36	27	52.9	47	1	US-07-819-893-9
37	27	52.9	450	1	US-08-194-338-5
38	27	52.9	330	1	US-08-118-270-21
39	27	52.9	214	1	US-08-318-492-4

Sequence 11, Appl  
Sequence 9, Appl  
Sequence 8, Appl  
Sequence 7, Appl  
Sequence 6, Appl  
Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-097-830E-10  
; Sequence 10, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-097-830E-10

Query Match 100.0%; Score 51; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
DB 1 KFFKFFKFF 9

RESULT 2  
US-08-456-112B-10  
; Sequence 10, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porto, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-10

Query Match 100.0%; Score 51; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 9  
Db 1 KFFKFFKFF 9

RESULT 3  
US-08-456-112B-40  
Sequence 40, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-40

Query Match 100.0%; Score 51; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 9  
Db 1 KFFKFFKFF 9

RESULT 4  
US-08-049-871-8  
Sequence 8, Application US/08049871  
Patent No. 5358933  
GENERAL INFORMATION:  
APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Synthetic Peptides for Detoxification  
TITLE OF INVENTION: Of Bacterial Endotoxins and for the  
TITLE OF INVENTION: Prevention and Treatment of Septic  
TITLE OF INVENTION: Shock  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/049,871  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/658,744  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-049-871-8

Query Match 64.7%; Score 33; DB 1; Length 11;  
Best Local Similarity 75.0%; Pred. No. 2.2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 8  
Db 2 KFLKFLKF 9



```

RESULT 5
US-07-819-893-8
; Sequence 8, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8998
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: circular
; US-07-819-893-8

Query Match 64.7%; Score 33; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKF 8
DB 2 KFLKFLKF 9

RESULT 6
US-08-280-397-8
; Sequence 8, Application US/08280397
; Patent No. 5589459
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8998
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acids
; TOPOLOGY: circular
; US-08-280-397-8

```

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,397
; FILING DATE: 07/26/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; US-08-280-397-8

Query Match 64.7%; Score 33; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 2.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKF 8
DB 2 KFLKFLKF 9

RESULT 7
US-08-097-830E-8
; Sequence 8, Application US/08097830E
; Patent No. 5652211
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Peptides For Neutralizing The
; TITLE OF INVENTION: Toxicity of Lipid A
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,830E
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acids
; TOPOLOGY: circular
; US-08-097-830E-8

```

TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-8

Query Match 64.7%; Score 33; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKF 8  
DB 1 KFLKFLKF 8

## RESULT 8

US-08-097-830E-20  
; Sequence 20, Application US/08097830E  
; Patent No. 5652211

## GENERAL INFORMATION:

APPLICANT: Porro, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TOXICITY OF LIPID A

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan, P.C.

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/097,830E

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

TOPOLOGY: circular

FEATURE:

OTHER INFORMATION: sulfide bond between Cys and Cys

US-08-097-830E-20

Query Match 64.7%; Score 33; DB 1; Length 12;  
Best Local Similarity 75.0%; Pred. No. 2.3;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKF 8  
DB 2 KFLKFLKF 9

## RESULT 9

US-08-456-112B-8

; Sequence 8, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

APPLICANT: Porro, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: LEADING EDGE 486

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,112B

FILING DATE: May 31, 1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Costigan, James V.

REGISTRATION NUMBER: 25,669

REFERENCE/DOCKET NUMBER: 576-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989

TELEFAX: (212) 302-8998

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

TOPOLOGY: circular

US-08-456-112B-8

Query Match 64.7%; Score 33; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFKFKFKF 8  
DB 1 KFLKFLKF 8

## RESULT 10

US-08-456-112B-20

; Sequence 20, Application US/08456112B

; Patent No. 5834430

; GENERAL INFORMATION:

APPLICANT: Porro, Massimo

TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hedman, Gibson & Costigan

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: LEADING EDGE 486

OPERATING SYSTEM: DOS

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,112B

FILING DATE: May 31, 1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-20

Query Match 64.7%; Score 33; DB 2; Length 12;  
Best Local Similarity 75.0%; Pred. No. 2.3;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKF 8  
||| ||  
Db 2 KFLKFLKF 9

RESULT 11  
US-08-456-112B-37  
Sequence 37, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: PORIO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-37

Query Match 64.7%; Score 33; DB 2; Length 11;  
Best Local Similarity 75.0%; Pred. No. 2.2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKF 8  
||| ||  
Db 2 KFLKFLKF 9

RESULT 12  
US-08-456-112B-41  
Sequence 41, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: PORIO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-41

Query Match 64.7%; Score 33; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKF 8  
||| ||  
Db 2 KFLKFLKF 9

RESULT 13  
US-08-323-170B-2  
Sequence 2, Application US/08323170B  
Patent No. 5733772  
GENERAL INFORMATION:  
APPLICANT: Williamson, Kim C.  
APPLICANT: Kaslow, David C.  
TITLE OF INVENTION: Cloning and Expression of Plasmodium  
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pfs230  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/323,170B  
FILING DATE: 13-OCT-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/010,409  
FILING DATE: 29-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Quine, Jonathan A.  
REGISTRATION NUMBER: P-41,261  
REFERENCE/DOCKET NUMBER: 015280-113100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3135 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-323-170B-2

Query Match 62.7%; Score 32; DB 1; Length 3135;  
Best Local Similarity 85.7%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKEKF 8  
| | | | |  
DB 993 FSKFKEF 999

## RESULT 14

US-07-867-106-3  
Sequence 3, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rls  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 887 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-867-106-3

Query Match 60.8%; Score 31; DB 1; Length 887;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KFFKEFF 6  
: | | | | |  
DB 9 RFFKEFF 14

## RESULT 15

US-08-426-428-2  
Sequence 2, Application US/08426428  
Patent No. 5637462  
GENERAL INFORMATION:  
APPLICANT: Coleman, Roger  
APPLICANT: Braxton, Scott M.  
APPLICANT: Seilhamer, Jeffrey J.  
TITLE OF INVENTION: NEW CATHEPSIN C HOMOLOG  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3330 Hillview Avenue  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/426,428  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PF-0032 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 463 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-426-428-2

Query Match 60.8%; Score 31; DB 1; Length 463;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKEFF 8  
: | | | | |  
DB 100 KWFKEFF 107

Search completed: September 7, 1999, 23:07:52  
Job time: 1750 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:22:59 ; Search time 105.14 Seconds  
(without alignments)  
3.430 Million cell updates/sec

Title: US-09-124-280A-40

Perfect score: 51

Sequence: 1 KFFKFFKFF 9

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	76.5	2010	2 B71616	phosphatase (acid
2	38	74.5	873	2 S53828	cytochrome-c oxida
3	37	72.5	241	2 S34500	hypothetical prote
4	36	70.6	988	2 S77211	isoleucine--trna 1
5	36	70.6	409	2 S74704	hypothetical prote
6	36	70.6	317	2 D64233	hypothetical prote
7	36	70.6	411	2 T01471	hypothetical prote
8	36	70.6	332	2 S31848	heat shock protein
9	36	70.6	136	2 S19479	hypothetical prote
10	36	70.6	1014	2 H71602	protein with DnaJ
11	35	68.6	174	2 H64007	hypothetical prote
12	35	68.6	615	2 S74277	probable membrane
13	35	68.6	615	2 S40824	probable membrane
14	35	68.6	824	2 S40937	hypothetical prote
15	34	66.7	1082	1 RNEGB	DNA-directed RNA p
16	34	66.7	193	2 S07734	NADH dehydrogenase
17	34	66.7	516	2 S34525	hypothetical prote
18	34	66.7	649	2 G71530	probable ATP synth
19	34	66.7	117	2 S66720	probable membrane
20	33	64.7	166	2 S10513	interferon gamma -
21	33	64.7	144	2 B30471	hypothetical prote
22	33	64.7	1307	2 S53412	probable membrane
23	33	64.7	192	2 H70138	cell division prot
24	33	64.7	569	2 A45624	trophozoite cystel
25	33	64.7	459	2 F70189	rod shape-determin
26	33	64.7	233	2 E64596	ABC transporter, A
27	33	64.7	376	2 F71828	hypothetical prote
28	33	64.7	382	2 E64686	myosin-like protei
29	33	64.7	233	2 E71950	probable ABC trans
30	33	64.7	455	2 JC5726	phosphate-phenyl 9
31	33	64.7	502	2 T00483	hypothetical prote
32	33	64.7	113	2 S07726	hypothetical prote
33	33	64.7	644	2 B70420	NADH dehydrogenase
34	32	62.7	495	1 S52641	heat shock transcr
35	32	62.7	3135	2 A48584	transmission block
36	32	62.7	612	2 S11408	hypothetical prote
37	32	62.7	312	2 S49612	transposase - Esch
38	32	62.7	598	2 H69011	hypothetical prote
39	32	62.7	501	2 A69415	conserved hypothet

40 32 62.7 169 2 S57138 hypothetical prote  
41 32 62.7 695 2 S61873 hypothetical prote  
42 32 62.7 491 2 S49790 probable membrane  
43 32 62.7 145 2 S63134 probable membrane  
44 32 62.7 258 2 S63194 probable membrane  
45 32 62.7 1802 2 G71616 hypothetical prote

ALIGNMENTS

RESULT 1  
B71616  
phosphatase (acid phosphatase family) PFB0380c - malaria parasite (Plasmodium falci  
C;Species: Plasmodium falciparum  
C;Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C;Accession: B71616  
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.  
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H  
Science 282, 1126-1132, 1998  
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A;Reference number: A71600  
A;Accession: B71616  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-2010 <GAR>  
A;Cross-references: GB:AE001391; GB:AE001362; NID:g3845168; PID:g3845169; TIGR:PFB038  
A;Experimental source: clone 3D7  
C;Genetics:  
A;Gene: PFB0380c

Query Match 76.5%; Score 39; DB 2; Length 2010;  
Best Local Similarity 87.5%; Pred No 29;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KFFKFFKF 8  
DB 801 KFFKLFKF 808  
RESULT 2  
S53828  
cytochrome-c oxidase (EC 1.9.3.1) chain 1 and chain 2 - Acanthamoeba castellanii mito  
C;Species: mitochondrion Acanthamoeba castellanii  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 29-Jan-1999  
C;Accession: S53828; S68552  
R;Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.  
J. Mol. Biol. 245, 522-537, 1995  
A;Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: c  
A;Reference number: S53825; MUID:95147275  
A;Accession: S53828  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-873 <BUR>  
A;Cross-references: GB:U12386; NID:g562028; PID:g562032  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
R;Loneragan, K.M.; Gray, M.W.  
J. Mol. Biol. 257, 1019-1030, 1996  
A;Title: Expression of a continuous open reading frame encoding subunits 1 and 2 of c  
A;Reference number: S68552; MUID:96192086  
A;Accession: S68552  
A;Molecule type: DNA  
A;Residues: 1-873 <LON>  
A;Cross-references: GB:U12386; NID:g562028; PID:g562032  
A;Experimental source: ATCC 30010  
C;Genetics:  
A;Gene: cox1/2  
A;Genome: mitochondrion  
A;Genetic code: SGC6  
C;Superfamily: Acanthamoeba cytochrome-c oxidase chain I/II; cytochrome-c oxidase cha  
C;Keywords: copper binding; electron transfer; heme; membrane-associated complex; mit  
F;23-470/Domain: cytochrome-c oxidase chain I homology <CO1>

F:568-854/Domain: cytochrome-c oxidase chain II homology <CO2>

Query Match 74.5%; Score 38; DB 2; Length 873;  
Best Local Similarity 87.5%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFKFF 9  
| | | | |  
Db 530 FFKSKFF 537

# RESULT 3

S34500  
hypothetical protein 241 (psbc 3' region) - Euglena gracilis chloroplast  
C:Species: chloroplast Euglena gracilis  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Apr-1995  
C:Accession: S34500; S34868  
R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielmann,  
submitted to the EMBL Data Library, January 1993  
A:Description: The complete sequence of the Euglena gracilis chloroplast genome (tentative)  
A:Reference number: S34494  
A:Accession: S34500  
A:Molecule type: DNA  
A:Residues: 1-241 <HAL1>  
A:Cross-references: EMBL:X70810  
R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielmann,  
Nucleic Acids Res. 21, 3537-3544, 1993  
A:Title: Complete sequence of Euglena gracilis chloroplast DNA.  
A:Reference number: S34862; MUID:93347989  
A:Accession: S34868  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-241 <HAL2>  
A:Cross-references: EMBL:X70810  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 72.5%; Score 37; DB 2; Length 241;  
Best Local Similarity 85.7%; Pred. No. 8.7;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 FFKFKFF 9  
| | | | |  
Db 98 FFKFKFF 104

# RESULT 4

S77211  
isoleucine--tRNA ligase (EC 6.1.1.5) - Synecocystis sp. (strain PCC 6803)  
N:Alternate names: isoleucyl-tRNA synthetase; protein sll1362  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S77211  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S77211  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-988 <KAN>  
A:Cross-references: EMBL:D90907; GB:AB001339; NID:g1652618; PID:d1018278; PID:g1652625  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: lles  
A:Start codon: GTG

C:Superfamily: isoleucine--tRNA ligase  
C:Keywords: ligase

Query Match 70.6%; Score 36; DB 2; Length 988;  
Best Local Similarity 85.7%; Pred. No. 47;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 FFKFKFF 9  
| | | | |  
Db 724 FFKFKFF 730

# RESULT 5

S74704  
hypothetical protein slr1302 - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S74704  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S74704  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-409 <KAN>  
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PID:d1017588; PID:g165192  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Start codon: GTG

Query Match 70.6%; Score 36; DB 2; Length 409;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
| | | | |  
Db 92 KFFKFFKFF 100

# RESULT 6

D64233  
hypothetical protein MG302 - Mycoplasma genitalium (SGC3)  
C:Species: Mycoplasma genitalium  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 10-Oct-1997  
C:Accession: D64233  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346  
A:Accession: D64233  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-317 <TIGR>  
A:Cross-references: GB:U39711; GB:L43967; NID:g1045997; PID:g1046001; TIGR:MG302  
A:Experimental source: strain G-37  
C:Genetics:  
A:Genetic code: SGC3

Query Match 70.6%; Score 36; DB 2; Length 317;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
| | | | |

Db 152 RFFKFLKLF 160

RESULT 7

T01471

hypothetical protein T24H24.4 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 24-Mar-1999

C:Accession: T01471

R:Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K.

submitted to the EMBL Data Library, August 1998

A:Description: The sequence of A. thaliana T24H24.

A:Reference number: 214333

A:Accession: T01471

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-411 <COU>

A:Cross-references: EMBL:AF075598; NID:g3293581; PID:g3377829

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 288/1

A:Note: T24H24.4

Query Match 70.6%; Score 36; DB 2; Length 411;

Best Local Similarity 75.0%; Pred. No. 21;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFKFKFF 9

||| ||| ||

Db 127 FFEFFTF 134

RESULT 8

S31848

heat shock protein HSP30 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YCR021c

C:Species: Saccharomyces cerevisiae

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 06-Feb-1998

C:Accession: S31848; S19432; S30781

R:Regnacq, M.; Boucherie, H.

Curr. Genet. 23, 435-442, 1993

A:Title: Isolation and sequence of HSP30, a yeast heat-shock gene coding for a hydrophob

A:Reference number: S31848; MUID:93306747

A:Accession: S31848

A:Molecule type: DNA

A:Residues: 1-263 <REG1>

A:Cross-references: EMBL:M93123

R:Feldmann, H.; Mannhaupt, G.; Vetter, I.

submitted to the Protein Sequence Database, March 1992

A:Reference number: S19429

A:Accession: S19432

A:Molecule type: DNA

A:Residues: 1-190, 'A', 192-332 <PEL>

A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264485; PID:g1907165; MIPS:YCR021c

R:Regnacq, M.; Boucherie, H.

submitted to the EMBL Data Library, January 1993

A:Reference number: S30781

A:Accession: S30781

A:Molecule type: DNA

A:Residues: 1-152, 'GY', 164-165, 167, 'A', 169, 'NSNRGL', 170-240, 'VFNOT', 270, 278-280, 'FMVFWI

A:Cross-references: EMBL:M93123

A:Note: the difference at the carboxyl end is due to a frameshift error

C:Genetics:

A:Gene: SGD:HSP30

A:Cross-references: SGD:S0000615; MIPS:YCR021c

A:Map position: 3R

C:Keywords: membrane protein

Query Match

Best Local Similarity 70.6%; Score 36; DB 2; Length 332;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFKFKFF 9

||| ||| ||

Db 76 FFEFFAF 83

RESULT 9

S19479

hypothetical protein YCR064c - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 14-Nov-1997

C:Accession: S19479

R:Antoine, G.; Benit, P.; Chanet, R.; Fabre, R.; Faye, G.; Fukuhara, H.; Mathieu, A.

submitted to the Protein Sequence Database, March 1992

A:Reference number: S19477

A:Accession: S19479

A:Molecule type: DNA

A:Residues: 1-136 <ANT>

A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264407; PID:g1907204; MIPS:YCR064

C:Genetics:

A:Map position: 3R

Query Match 70.6%; Score 36; DB 2; Length 136;

Best Local Similarity 66.7%; Pred. No. 7.4;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFFKFFFF 9

||| ||| ||

Db 88 KFFFFQYF 96

RESULT 10

H71602

protein with DnaJ domain (RESA-like) PFB0920w - malaria parasite (Plasmodium falcipar

C:Species: Plasmodium falciparum

C>Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 16-Dec-1998

C:Accession: H71602

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

.; Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.

Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600

A:Accession: H71602

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1014 <GAR>

A:Cross-references: GB:AE001425; GB:AE001362; NID:g3845310; PID:g3845311; TIGR:PFB092

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0920w

C:Superfamily: dnaJ amino-terminal homology

F:677-742/Domain: dnaJ amino-terminal homology <DNJ>

Query Match

Best Local Similarity 70.6%; Score 36; DB 2; Length 1014;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFFKFFFF 9

||| ||| ||

Db 2 KYFKFKYF 10

RESULT 11

H64007

hypothetical protein HI0453 - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 30-Jun-1998

C:Accession: H64007

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630  
 A:Accession: H64007  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-174 <TIGR>  
 A:Cross-references: GB:U32728; GB:I42023; NID:g1573425; PID:g1573438; TIGR:HI0453  
 C:Superfamily: conserved hypothetical protein HI0453

Query Match 68.6%; Score 35; DB 2; Length 174;  
 Best Local Similarity 75.0%; Pred. No. 14;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKF 8  
 |||||  
 DB 3 KFFKYFLF 10

RESULT 12  
 S74277  
 Probable membrane protein YCL073c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein YCL070c; hypothetical protein YCL071c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 04-Sep-1998  
 A:Accession: S74277; S69990; S19402; S19403; S19405  
 R:Voet, M.; Volckaert, G.  
 A:Reference number: S74277  
 A:Reference number: S74277  
 A:Accession: S74277  
 A:Molecule type: DNA  
 A:Residues: 1-615 <VOE>  
 A:Cross-references: EMBL:X59720; NID:g1907116; PID:g309031; PID:g1907117; MIPS:YCL073c  
 A:Note: this is a revision to the sequence from reference S19400  
 R:Gromadka, R.; Gora, M.; Zielenkiewicz, U.; Slonimski, P.P.; Rytka, J.  
 Yeast 12, 583-591, 1996  
 A:Title: Subtelomeric duplications in Saccharomyces cerevisiae chromosomes III and XI:  
 A:Reference number: S69990; MUID:96367599  
 A:Accession: S69990  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-465, 'I', 467-526, 'N', 528-533, 'S', 535-615 <GRO>  
 A:Cross-references: EMBL:X59720  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
 R:van der Aart, O.J.M.; Steensma, H.Y.  
 submitted to the Protein Sequence Database, March 1992  
 A:Reference number: S19400  
 A:Accession: S19402  
 A:Molecule type: DNA  
 A:Residues: 1-214, 'G', 216-270, 'L', 272-326, 'PNLPNLLF', 335-349, 'VLL', 353, 'SI', 354, 'SP', 357  
 A:Cross-references: EMBL:X59720  
 A:Note: this was assumed to be protein YCL070c  
 A:Note: this sequence has been revised in reference S74277  
 A:Accession: S19403  
 A:Molecule type: DNA  
 A:Residues: 374-484, 'ALLSVPLFVPGKPKLTNSTLTETIH', 497, 'QK' <NAW>  
 A:Cross-references: EMBL:X59720  
 A:Note: this was assumed to be protein YCL071c  
 A:Note: this sequence has been revised in reference S74277  
 A:Accession: S19405  
 A:Molecule type: DNA  
 A:Residues: 501-586, 'VG', 589-590, 'T', 592-604, 'R', 606-609, 'N', 611, 'KL', 614-615 <AAF>  
 A:Cross-references: EMBL:X59720; MIPS:YCL073c  
 A:Note: this sequence has been revised in reference S74277  
 A:Note: this was assumed to be the complete sequence of protein YCL073c  
 C:Genetics:  
 A:Map position: 3L  
 A:Note: YCL073c  
 C:Superfamily: conserved probable membrane protein YCL073c  
 C:Keywords: transmembrane protein

F:60-76/Domain: transmembrane #status predicted <TM1>  
 F:125-144/Domain: transmembrane #status predicted <TM2>  
 F:220-236/Domain: transmembrane #status predicted <TM3>  
 F:276-292/Domain: transmembrane #status predicted <TM4>  
 F:308-324/Domain: transmembrane #status predicted <TM5>  
 F:356-372/Domain: transmembrane #status predicted <TM6>  
 F:411-427/Domain: transmembrane #status predicted <TM7>  
 F:445-461/Domain: transmembrane #status predicted <TM8>  
 F:549-565/Domain: transmembrane #status predicted <TM9>

Query Match 68.6%; Score 35; DB 2; Length 615;  
 Best Local Similarity 75.0%; Pred. No. 44;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 |||||  
 DB 353 FFNFFTFF 360

RESULT 13  
 S40524  
 Probable membrane protein YCL073c homolog YKR106w - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 04-Sep-1998  
 C:Accession: S40524  
 R:Gaillon, L.; Dujon, B.  
 submitted to the Protein Sequence Database, March 1994  
 A:Reference number: S38175  
 A:Accession: S40624  
 A:Molecule type: DNA  
 A:Residues: 1-615 <GAI>  
 A:Cross-references: EMBL:Z28202; NID:g486616; PID:g486617; MIPS:YKR106w  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Map position: 11R  
 C:Superfamily: conserved probable membrane protein YCL073c  
 C:Keywords: transmembrane protein

Query Match 68.6%; Score 35; DB 2; Length 615;  
 Best Local Similarity 75.0%; Pred. No. 44;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 |||||  
 DB 353 FFNFFTFF 360

RESULT 14  
 S40937  
 Hypothetical protein ZK632.5 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Sep-1997  
 C:Accession: S40937  
 R:Berk, M.  
 submitted to the EMBL Data Library, February 1993  
 A:Reference number: S40933  
 A:Accession: S40937  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-824 <BER>  
 A:Cross-references: EMBL:Z22181; NID:g297998; PID:g298003  
 C:Genetics:  
 A:Introns: 43/3; 74/1; 112/3; 159/1; 318/2; 363/3; 400/2; 575/2; 607/1; 696/3; 747/3;

Query Match 68.6%; Score 35; DB 2; Length 824;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 |||||



Db 553 KFFFLKLF 561

## RESULT 15

RNEGB

DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Euglena gracilis chloroplast

C:Species: chloroplast Euglena gracilis

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 05-Sep-1997

C:Accession: S09210

R:Yepiz-Plascencia, G.M.; Radebaugh, C.A.; Hallick, R.B.

Nucleic Acids Res. 18, 1869-1878, 1990

A:Title: The Euglena gracilis chloroplast rpoB gene. Novel gene organization and transcr

A:Reference number: S09210; MUID:90245579

A:Accession: S09210

A:Molecule type: DNA

A:Residues: 1-1082 &lt;VEP&gt;

A:Cross-references: EMBL:X17191; NID:g11501; PID:g311709

A:Note: the authors translated the codon CGT for residue 132 as Gly

C:Genetics:

A:Gene: rpoB

A:Genome: chloroplast

A:Introns: 13/3; 89/2; 117/1; 154/3; 227/3; 241/1; 288/2; 732/3

C:Superfamily: DNA-directed RNA polymerase beta chain

C:Keywords: chloroplast; nucleotidyltransferase; transcription

## Query Match

66.7%; Score 34; DB 1; Length 1082;

Best Local Similarity 55.6%; Pred. No. 1.1e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFFKFFKFF 9

:|:|:|

Db 210 EFFEYKTF 218

Search completed: September 7, 1999, 23:23:00

Job time: 967 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:59:19 ; Search time 71.09 Seconds  
(without alignments)  
3.579 Million cell updates/sec

Title: US-09-124-280A-40

Perfect score: 51 KFFKFFKFF 9

Sequence: 1 KFFKFFKFF 9

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	72.5	634	1	YCX3_EUGGR
2	36	70.6	332	1	HS30_YEAST
3	36	70.6	988	1	SVI_SYNY3
4	36	70.6	317	1	X302_MYCGE
5	35	70.6	136	1	YCV4_YEAST
6	35	68.6	616	1	SRE2_CAEEL
7	35	68.6	174	1	Y453_HAEIN
8	35	68.6	615	1	YCHO_YEAST
9	35	68.6	615	1	YK86_YEAST
10	35	68.6	824	1	YOT5_CAEEL
11	34	66.7	193	1	NU2M_PASTE
12	34	66.7	1082	1	RPOB_EUGGR
13	34	66.7	516	1	YCX4_EUGGR
14	34	66.7	504	1	Y093_CAEEL
15	33	64.7	569	1	CVSP_PLAFA
16	33	64.7	166	1	ING_PIG
17	33	64.7	113	1	YK01_PASTE
18	32	62.7	483	1	HSF1_ARATH
19	32	62.7	3135	1	S230_PLAFO
20	32	62.7	491	1	YIJO_YEAST
21	32	62.7	169	1	YJ85_YEAST
22	32	62.7	156	1	YM04_PASTE
23	32	62.7	145	1	YNR9_YEAST
24	32	62.7	258	1	YMW8_YEAST
25	31	60.8	463	1	CATC_HUMAN
26	31	60.8	462	1	CATC_MOUSE
27	31	60.8	462	1	CATC_RAT
28	31	60.8	288	1	CDSA_HAEIN
29	31	60.8	574	1	DPOL_AQUAE
30	31	60.8	990	1	DPOL_NPVCF
31	31	60.8	254	1	FLIP_BORBU
32	31	60.8	416	1	LACY_CITFR
33	31	60.8	497	1	NU4M_ACACA
34	31	60.8	287	1	NU4M_MYTEO
35	31	60.8	578	1	PSP2_YEAST
36	31	60.8	1129	1	YB95_YEAST
37	31	60.8	1738	1	YCF1_EPIVI
38	31	60.8	796	1	YEZ9_YEAST
39	31	60.8	497	1	YFC7_YEAST
40	31	60.8	616	1	YGO2_YEAST
41	31	60.8	639	1	YHR7_YEAST
42	31	60.8	371	1	YHSC_CLOAB
43	31	60.8	407	1	YKV9_YEAST

ALIGNMENTS

RESULT 1

YCX3\_EUGGR  
ID YCX3\_EUGGR STANDARD; PRT; 634 AA.  
AC P31916; P31917;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 78.0 KD PROTEIN IN PSBC INTRON 2 (ORF635).  
OS EUGLENA GRACILIS.  
OG CHLOROPLAST.  
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Z;  
RX MEDLINE: 93347989.  
RA HALLICK R.B., HONG L., DRAGER R.G., FAVREAU M.R., MONFORT A.,  
RA ORSAT B., SPIELMANN A., STUTZ E.;  
RT \*Complete sequence of Euglena gracilis chloroplast DNA.\*;  
RL NUCLEIC ACIDS RES. 21:3537-3544(1993).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; Z11874; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X70810; G415736; -;  
DR PIR; S34499; S34499.  
DR PIR; S34500; S34500.  
KW CHLOROPLAST; HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 634 AA; 78049 MW; 9EE16D0D CRC32;

Query Match 72.5% Score 37; DB 1; Length 634;  
Best Local Similarity 85.7%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KFFKFF 9

Db 491 FRFFKFF 497

RESULT 2

HS30\_YEAST  
ID HS30\_YEAST STANDARD; PRT; 332 AA.  
AC P25619; Q04556;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE 30 KD HEAT SHOCK PROTEIN.  
GN HSP30 OR YCR021C OR YCR21C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SKQ2N;  
RX MEDLINE: 93306747.  
RA REGNAQ M., BOUCHERIE H.;  
RT \*Isolation and sequence of HSP30, a yeast heat-shock gene coding for  
RT a hydrophobic membrane protein.\*;  
RL CURR. GENET. 23:435-442(1993).  
RN [2]

SEQUENCE FROM N.A.  
 RA FELDMANN H., MANNHAUPT G., VETTER I.,  
 RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: PROBABLY COOPERATES WITH OTHER HEAT SHOCK PROTEINS IN  
 CC IN THE TRANSLOCATION OF POLYPEPTIDES THROUGH MEMBRANES. IT MAY  
 CC COUNTERACT THE ALTERING EFFECT OF HEAT SHOCK ON THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING THE ENTRY INTO STATIONARY  
 CC PHASE RESULTING FROM GLUCOSE LIMITATION.  
 CC -!- SIMILARITY: TO YEAST YRO2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M93123; G405821; -  
 CC EMBL: X59720; E264485; -  
 CC PIR: S30781; S30781.  
 CC PIR: S19432; S19432.  
 CC SGD: L0000818; HSP30.  
 CC HEAT SHOCK; TRANSMEMBRANE.  
 CC FT TRANSMEM 35 55 POTENTIAL.  
 CC FT TRANSMEM 66 86 POTENTIAL.  
 CC FT TRANSMEM 122 142 POTENTIAL.  
 CC FT TRANSMEM 158 178 POTENTIAL.  
 CC FT TRANSMEM 182 202 POTENTIAL.  
 CC FT TRANSMEM 216 236 POTENTIAL.  
 CC FT TRANSMEM 249 269 POTENTIAL.  
 CC FT DOMAIN 290 332 GLU-RICH (ACIDIC).  
 CC SEQUENCE 332 AA; 37044 MW; FCD9C61B CRC32;

Query Match 70.6%; Score 36; DB 1; Length 332;  
 Best Local Similarity 75.0%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FKFKEFF 9  
 Db 76 FFEFFAFF 83

RESULT 3  
 SYL\_SYN3  
 ID SYL\_SYN3 STANDARD; PRT: 988 AA.  
 AC P73505;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ISOLEUCYL-TRNA SYNTHETASE (EC 6.1.1.5) (ISOLEUCINE--TRNA LIGASE)  
 DE (ILERS).  
 GN ILES OR SLL1362.  
 OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).  
 OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 97061201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K.,  
 RA OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A.,  
 RA TAMADA M., YASUDA M., TABATA S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA RES. 3:109-136(1996).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-ISOLEUCINE + TRNA(ILE) = AMP +  
 CC PYROPHOSPHATE + L-ISOLEUCYL-TRNA(ILE).  
 CC -!- COFACTOR: BINDS A ZINC ION (BY SIMILARITY).

CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D90907; G1652625; -  
 CC DR PROSITE; PS00178; AA-TRNA\_LIGASE\_I; 1.  
 CC DR PFAM: PF00133; TRNA-SYNT\_1; 1.  
 CC KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING;  
 CC METAL-BINDING; ZINC. 70  
 CC FT SIMILAR 60 "HIGH" REGION.  
 CC FT SIMILAR 611 615 "RMSKS" REGION.  
 CC FT BINDING 614 614 ATP (BY SIMILARITY).  
 CC SEQUENCE 988 AA; 112345 MW; C89678C2 CRC32;  
 CC  
 CC Query Match 70.6%; Score 36; DB 1; Length 988;  
 CC Best Local Similarity 85.7%; Pred. No. 34;  
 CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 3 FKFKEFF 9  
 Db 724 FFFKFF 730  
 CC  
 CC RESULT 4  
 Y302\_MYCGE  
 ID Y302\_MYCGE STANDARD; PRT: 317 AA.  
 AC P47544;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN MG302.  
 GN MG302.  
 OS MYCOPLASMA GENITALIUM.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;  
 OC MYCOPLASMATACEAE; MYCOPLASMA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 33530 / G-37;  
 RX MEDLINE: 96026346.  
 RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,  
 RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,  
 RA FRITZCHMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,  
 RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,  
 RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,  
 RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium";  
 RT SCIENCE 270:397-403(1995).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U39711; G1046001; -  
 CC TIGR: MG302; -  
 CC KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 CC FT TRANSMEM 18 38 POTENTIAL.  
 CC FT TRANSMEM 58 78 POTENTIAL.  
 CC FT TRANSMEM 92 112 POTENTIAL.  
 CC FT TRANSMEM 130 150 POTENTIAL.

FT TRANSMEM 159 179 POTENTIAL.  
 FT TRANSMEM 202 222 POTENTIAL.  
 FT TRANSMEM 252 272 POTENTIAL.  
 SQ SEQUENCE 317 AA: 36667 MW; EF5F2980 CRC32;

Query Match 70.6%; Score 36; DB 1; Length 317;

Best Local Similarity 66.7%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9

Db 152 REFFKFLKF 160

RESULT 5

YCV4\_YEAST STANDARD; PRT: 136 AA.

ID YCV4\_YEAST STANDARD; PRT: 136 AA.

AC P25640;

DT 01-MAY-1992 (REL. 22, CREATED)

DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 15.8 KD PROTEIN IN SUP61-RAD18 INTERGENIC REGION.

GN YCR064C OR YCR64C.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

OC SACCHAROMYCETACEAE; SACCHAROMYCES.

RN [1]

RP SEQUENCE FROM N.A.

RA ANTOINE G., BENIT P., CHANET R., FABRE R., FAYE G., FUKUHARA H.,

RA MATHIEU A., SOR F.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X59720; E264407; -

DR PIR: S19479; S19479.

KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 136 AA: 15776 MW; 8FACF40B CRC32;

Query Match 70.6%; Score 36; DB 1; Length 136;

Best Local Similarity 66.7%; Pred. No. 5.5;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9

Db 88 KFFFFFOYF 96

RESULT 6

SRE2\_CAEEL

ID SRE2\_CAEEL STANDARD; PRT: 616 AA.

AC Q09273;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE SRE-2 PROTEIN.

GN SRE-2 OR C41C4.2.

OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTERA; RHABDITIA; RHABDITIDA;

OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA BURTON J.;

RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -|- SIMILARITY: BELONGS TO THE C.ELEGANS RECEPTOR-LIKE PROTEIN SRE  
 CC FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: Z48045; E1344654; -

DR WORMPEP: C41C4.2; CE01517.

KW TRANSMEMBRANE; MULTIGENE FAMILY.

FT TRANSMEM 7 27 POTENTIAL.

FT TRANSMEM 129 149 POTENTIAL.

FT TRANSMEM 179 199 POTENTIAL.

FT TRANSMEM 269 289 POTENTIAL.

FT TRANSMEM 303 323 POTENTIAL.

FT TRANSMEM 396 416 POTENTIAL.

FT TRANSMEM 426 446 POTENTIAL.

FT TRANSMEM 482 502 POTENTIAL.

FT TRANSMEM 522 542 POTENTIAL.

SQ SEQUENCE 616 AA: 70480 MW; 2FC75724 CRC32;

Query Match 68.6%; Score 35; DB 1; Length 616;

Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9

Db 4 FFKFFSIF 11

RESULT 7

Y453\_HAEIN

ID Y453\_HAEIN STANDARD; PRT: 174 AA.

AC P43999;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL PROTEIN HI0453.

GN HI0453.

OS HAEMOPHILUS INFLUENZAE.

OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;

OC HAEMOPHILUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-RD / KW20;

EX MEDLINE; 95350630.

RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,

RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,

RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,

RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,

RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,

RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,

RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEORGHEN N.S.M.,

RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,

RA VENTER J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus

influenzae Rd."

RL SCIENCE 269:496-512(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U32728; G1573438; -.  
 KW TIGR; H10453; -.  
 FT TRANSMEM 8 28 POTENTIAL.  
 FT TRANSMEM 146 166 POTENTIAL.  
 SQ SEQUENCE 174 AA; 20416 MW; 23634093 CRC32;

Query Match 68.6%; Score 35; DB 1; Length 174;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 KFKFKFF 8  
 |||||  
 Db 3 KFKFYLF 10

## RESULT 8

YCHO\_YEAST STANDARD; PRT; 615 AA.  
 AC P25596; P25597; P25599;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 68.9 KD PROTEIN IN HML 5' REGION.  
 GN YCLO73C OR YCLO70C/71C/73C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RQ SEQUENCE FROM N.A.

RA VAN DER AART Q.J.M., STERNESMA H.Y.;  
 RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP REVISIONS.

RA GROMADKA R.;  
 RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE YCLO70C/YHLO47C/YKRI06W FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X59720; E264591; -.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 59 79 POTENTIAL.  
 FT TRANSMEM 121 141 POTENTIAL.  
 FT TRANSMEM 153 173 POTENTIAL.  
 FT TRANSMEM 187 207 POTENTIAL.  
 FT TRANSMEM 217 237 POTENTIAL.  
 FT TRANSMEM 276 296 POTENTIAL.  
 FT TRANSMEM 308 328 POTENTIAL.  
 FT TRANSMEM 344 364 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 409 429 POTENTIAL.  
 FT TRANSMEM 441 461 POTENTIAL.  
 FT TRANSMEM 472 492 POTENTIAL.  
 FT TRANSMEM 548 568 POTENTIAL.  
 SQ SEQUENCE 615 AA; 68898 MW; 0BD5B37 CRC32;

Query Match 68.6%; Score 35; DB 1; Length 615;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFF 9  
 |||||  
 Db 353 FNFNFTFF 360

## RESULT 9

YK86\_YEAST STANDARD; PRT; 615 AA.  
 AC P36173;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 69.0 KD PROTEIN IN SIR1 3' REGION.  
 GN YKRI06W.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RQ SEQUENCE FROM N.A.

RA GAILLON L., DUJON B.;

RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE YCLO70C/YHLO47C/YKRI06W FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z28202; G486617; -.  
 DR PIR; S40624; S40624.

KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 59 79 POTENTIAL.  
 FT TRANSMEM 121 141 POTENTIAL.  
 FT TRANSMEM 153 173 POTENTIAL.  
 FT TRANSMEM 217 237 POTENTIAL.  
 FT TRANSMEM 276 296 POTENTIAL.  
 FT TRANSMEM 308 328 POTENTIAL.  
 FT TRANSMEM 344 364 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 408 428 POTENTIAL.  
 FT TRANSMEM 441 461 POTENTIAL.  
 FT TRANSMEM 472 492 POTENTIAL.  
 FT TRANSMEM 549 569 POTENTIAL.  
 SQ SEQUENCE 615 AA; 68954 MW; A8423CFB CRC32;

Query Match 68.6%; Score 35; DB 1; Length 615;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFKFF 9  
 |||||  
 Db 353 FNFNFTFF 360

## RESULT 10

YOT5\_CAEEL STANDARD; PRT; 824 AA.  
 AC P34651;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 95.7 KD PROTEIN ZK632.5 IN CHROMOSOME III.  
 GN ZK632.5.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RQ SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718;  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,  
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,  
 RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISSTER N.,  
 RA LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA STMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN K.,  
 RA SULSTON J., THIERREY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,  
 RA WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,  
 RA WORLDNAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL NATURE 368:32-38(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: 22181; E1351189; -  
 DR EMBL: 22181; E1351189; -  
 DR PIR: S40937; S40937;  
 DR WORMPEP: ZK632.5; CE00422.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 824 AA; 95726 MW; F3506E80 CRC32;

Query Match 68.68; Score 35; DB 1; Length 824;  
 Best Local Similarity 66.7%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 |||..|||  
 DB 553 KFFFLKLF 561

RESULT 11  
 NU2M\_PARTE  
 ID NU2M\_PARTE STANDARD; PRT; 193 AA.  
 AC P15577;  
 DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).  
 GN ND2 OR NDH2  
 OS PARAMECIUM TETRAURELIA.  
 OG MITOCHONDRION.  
 OC EUKARYOTA; ALVEOLATA; CILIOPHORA; NASSOPHOREA; PENICULIDA; PARAMECIUM.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-STOCK 51;  
 CC MEDLINE; 90174913.  
 RA PRITCHARD A.E., SEILHAMER J.J., MAHALINGAM R., SABLE C.L.,  
 RA VENUTI S.E., CUMMINGS D.J.;  
 RT "Nucleotide sequence of the mitochondrial genome of Paramecium.";  
 RL NUCLEIC ACIDS RES. 18:173-180(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-STOCK 51;  
 RX MEDLINE; 89357489.  
 RA PRITCHARD A.E., VENUTI S.E., GHALAMBOR M.A., SABLE C.L.,  
 RA CUMMINGS D.J.;  
 RT "An unusual region of Paramecium mitochondrial DNA containing  
 RT chloroplast-like genes.";  
 RL GENE 78:121-134(1989).  
 CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
 CC INNER MEMBRANE  
 CC -!- SIMILARITY: VERY DIFFERENT FROM OTHER ND2 SUBUNITS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M26930; G529481; -  
 DR EMBL: M26930; G529481; -  
 DR PIR: S07734; S07734;  
 DR PIR: JS0233; JS0233;  
 KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.  
 SQ SEQUENCE 193 AA; 23181 MW; 47D23D79 CRC32;

Query Match 66.7%; Score 34; DB 1; Length 193;  
 Best Local Similarity 75.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 |||||  
 DB 166 FFFFFKFF 173

RESULT 12  
 RPOB\_EUGGR  
 ID RPOB\_EUGGR STANDARD; PRT; 1082 AA.  
 AC P23579;  
 DT 01-NOV-1991 (REL. 20, CREATED)  
 DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6).  
 GN RPOB.  
 OS EUGLENA GRACILIS.  
 OG CHLOROPLAST.  
 OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-2;  
 CC MEDLINE; 90245579.  
 RA YEP12-PLASCENCIA G.M., RADEBAUGH C.A., HALLICK R.B.;  
 RT "The Euglena gracilis chloroplast rpoB gene. Novel gene organization  
 RT and transcription of the RNA polymerase subunit operon.";  
 RL NUCLEIC ACIDS RES. 18:1869-1878(1990).  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N).  
 CC -!- SUBUNIT: IN CHLOROPLAST THE RNA POLYMERASE IS COMPOSED OF FOUR  
 CC SUBUNITS: ALPHA, BETA, BETA', AND BETA".  
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: X17191; G311709; -  
 DR EMBL: X17191; G311709; -  
 DR PIR: S09210; RNEGB.  
 DR PROSITE; PS01166; RNA\_POL\_BETA; 1.  
 DR PFAM; PF00562; RNA\_POL\_B; 1.  
 DR MENDEL; 4737; EUGGR; rpoB; 1.  
 KW TRANSCRIPTION; DNA-DIRECTED RNA POLYMERASE; CHLOROPLAST.  
 SQ SEQUENCE 1082 AA; 124531 MW; 7510CC6F CRC32;

Query Match 66.7%; Score 34; DB 1; Length 1082;  
 Best Local Similarity 55.6%; Pred. No. 79;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKF 9  
 :||:|:|  
 Db 210 EFFEYKIF 218

## RESULT 13

YCX4\_EUGGR  
 ID YCX4\_EUGGR STANDARD; PRT; 516 AA.  
 AC P30397;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 64.3 KD PROTEIN IN RPS3 3'REGION (ORF516).  
 OS EUGLENA GRACILIS.  
 OG CHLOROPLAST.  
 CC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Z;  
 RX MEDLINE: 93347989.  
 RA HALLICK R.B., HONG L., DRAGER R.G., FAVREAU M.R., MONFORT A.,  
 RA ORSAT B., SPIELMANN A., STUTZ E.;  
 RT "Complete sequence of Euglena gracilis chloroplast DNA.";  
 RL NUCLEIC ACIDS RES. 21:3537-3544(1993).  
 CC -!- SIMILARITY: TO GROUP II INTRON MATURASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X70810; E1204055; -.  
 DR EMBL; Z11874; E1204062; -.  
 DR PIR; S26088; S26088.  
 DR PIR; S34525; S34525.  
 KW CHLOROPLAST; HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 516 AA; 64333 MW; CE19DE9F CRC32;

Query Match 66.7%; Score 34; DB 1; Length 516;  
 Best Local Similarity 75.0%; Pred. No. 40;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKF 8  
 ||:||||  
 Db 244 KYFFFKF 251

## RESULT 14

Y093\_CAEEL  
 ID Y093\_CAEEL STANDARD; PRT; 504 AA.  
 AC P41843;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 58.5 KD PROTEIN T20B12.3 IN CHROMOSOME III.  
 GN T20B12.3.  
 OS CAENORHABDITIS ELEGANS.  
 CC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
 CC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2.  
 RA WATERSTON R.;  
 RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U10401; G500716; -.  
 DR WORMPEP; T20B12.3; CE01409.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 504 AA; 58514 MW; BCCDDEB9 CRC32;

Query Match 66.7%; Score 34; DB 1; Length 504;  
 Best Local Similarity 66.7%; Pred. No. 39;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KFFKFFKF 9  
 |:||| |  
 Db 66 KYFKTKFF 74

## RESULT 15

CYSP\_PLAFA  
 ID CYSP\_PLAFA STANDARD; PRT; 569 AA.  
 AC P25805;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE THROPHOZOITE CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-) (TCP).  
 OS PLASMODIUM FALCIPARUM.  
 CC EUKARYOTA; ALVEOLATA; APICOMPLEXA; HAEMOSPORIDA; PLASMODIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92228005.  
 RA ROSENTHAL P.J., NELSON R.G.;  
 RT "Isolation and characterization of a cysteine proteinase gene of  
 RT Plasmodium falciparum.";  
 RL MOL. BIOCHEM. PARASITOL. 51:143-152(1992).  
 CC -!- FUNCTION: PROBABLY DEGRADES ERYTHROCYTE HEMOGLOBIN.  
 CC -!- DEVELOPMENTAL STAGE: THROPHOZOITE.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE  
 CC PAPAIA FAMILY OF THIOL PROTEASES.  
 CC -!- SIMILARITY: STRONGEST SIMILARITY TO CATHEPSIN L.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M81341; G160248; -.  
 DR PIR; A45624; A45624.  
 DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
 DR PFAM; PF00112; Cys-protease; 2.  
 DR HSP; P07858; ICSB.  
 KW HYDROLASE; THIOL PROTEASE; ZYMOMEN; GLYCOPROTEIN; SIGNAL.  
 RN SIGNAL 1 ?  
 FT PROPEP ? 332  
 FT CHAIN 333 569  
 FT DOMAIN 64 70  
 FT ACT\_SITE 357 357  
 FT ACT\_SITE 488 488  
 FT ACT\_SITE 533 533  
 FT DISULFID 354 395  
 FT CARBOHYD 58 58  
 FT CARBOHYD 98 98  
 FT CARBOHYD 121 121  
 FT CARBOHYD 127 127  
 FT CARBOHYD 479 479



FT CARBOHYD 487 487 POTENTIAL.  
SQ SEQUENCE 569 AA; 66880 MW; 9B8F0096 CRC32;

Query Match 64.7%; Score 33; DB 1; Length 569;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KFFKFFK 7  
      |||||  
Db 224 KFFKFMK 230

Search completed: September 7, 1999, 23:59:20  
Job time: 532 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:50 ; Search time 160.2 Seconds  
(without alignments)  
3.458 Million cell updates/sec

Title: us-09-124-280a-40

Perfect score: 51

Sequence: 1 KFFKFFKFF 9

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_protist.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	78.4	980	5	O17244 caenorhabdi
2	39	76.5	2010	5	O96171 plasmodium
3	38	74.5	303	2	O44421 anaerocellu
4	38	74.5	873	8	O37370 acanthamoeb
5	37	72.5	169	5	P91319 caenorhabdi
6	37	72.5	592	5	O17265 caenorhabdi
7	37	72.5	548	5	O44959 caenorhabdi
8	36	70.6	409	2	P72840 synchocyst
9	36	70.6	414	2	O92N10 staphylococ
10	36	70.6	709	3	O59824 schizosacch
11	36	70.6	430	5	O01449 caenorhabdi
12	36	70.6	619	5	O18207 caenorhabdi
13	36	70.6	1014	5	O96276 plasmodium
14	36	70.6	738	5	O97037 hydra magni
15	35	70.6	411	10	O81438 arabidopsis
16	35	68.6	615	3	P87002 saccharomyc
17	35	68.6	261	5	O76614 caenorhabdi
18	35	68.6	339	5	P91891 drosophila
19	34	66.7	427	2	O33751 synchococc
20	34	66.7	825	2	O59319 caldocellum
21	34	66.7	649	2	O84307 chlamydia t
22	34	66.7	117	3	O08205 saccharomyc
23	34	66.7	328	5	O21396 caenorhabdi
24	34	66.7	208	5	O18525 caenorhabdi
25	34	66.7	807	5	O17379 caenorhabdi
26	34	66.7	468	5	O44439 caenorhabdi
27	34	66.7	4981	5	O77372 plasmodium
28	34	66.7	495	5	O19453 caenorhabdi
29	33	64.7	455	2	P72577 streptococc

30	33	64.7	382	2	O25891 helicobacte
31	33	64.7	192	2	O51293 borrelia bu
32	33	64.7	459	2	O51661 borrelia bu
33	33	64.7	644	2	O67389 aquifex aeo
34	33	64.7	233	2	O25334 helicobacte
35	33	64.7	233	2	O92MCI helicobacte
36	33	64.7	376	2	O92JQ1 helicobacte
37	33	64.7	1307	3	O06412 saccharomyc
38	33	64.7	381	5	O94198 caenorhabdi
39	33	64.7	395	5	O17129 caenorhabdi
40	33	64.7	261	5	O44194 caenorhabdi
41	33	64.7	566	5	O45205 caenorhabdi
42	33	64.7	103	5	O97315 plasmodium
43	33	64.7	139	8	O48241 denhyus som
44	33	64.7	410	8	O47572 onchocerca
45	33	64.7	60	8	O20195 chlorella v

ALIGNMENTS

RESULT 1  
O17244  
ID O17244 PRELIMINARY; PRT; 980 AA.  
AC O17244;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE F53G2.6 PROTEIN.  
GN F53G2.6  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
ON Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE; 94150718;  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COFSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.;  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX BECKER M., WOHLDMANN P., BIEWALD T.;  
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX WATERSTON R.;  
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF025464; AAB71016.1; -;  
SQ SEQUENCE 980 AA; 111010 MW; 4A69D72C CRC32;

Query Match 78.4%; Score 40; DB 5; Length 980;  
Best Local Similarity 87.5%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFFKFFKFF 9  
DB 45 FFFKFFKFF 52

```

RESULT 2
O96171 ID O96171 PRELIMINARY; PRT; 2010 AA.
AC O96171;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE PHOSPHATASE (ACID PHOSPHATASE FAMILY).
GN PF0380C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RN [1]
RP SEQUENCE FROM N.A.
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001391; AAC71865.1; -.
SQ SEQUENCE 2010 AA; 241791 MW; 46C82357 CRC32;

Query Match 76.5%; Score 39; DB 5; Length 2010;
Best Local Similarity 87.5%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KFFKFFKF 8
Db 801 KFFKLFKF 808

RESULT 3
Q44421 ID Q44421 PRELIMINARY; PRT; 303 AA.
AC Q44421;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE SUGAR-BINDING TRANSPORT PROTEIN.
OS Anaerococcus thermophilum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Anaerococcus group; Anaerococcus.
RN [1]
RP SEQUENCE FROM N.A.
RA ZVERLOV V., ASCHERL G., VELIKODVORSKAYA G., BRONNENMEIER K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 277855; CAB01403.1; -.
DR PFAM; PF00528; BPD_transp; 1.
SQ SEQUENCE 303 AA; 35132 MW; CF756347 CRC32;

Query Match 74.5%; Score 38; DB 2; Length 303;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KFFKFFKF 9
Db 114 KFFTFGGF 122

RESULT 4
Q37370 ID Q37370 PRELIMINARY; PRT; 873 AA.
AC Q37370;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-WAY-1999 (TREMBlrel. 10, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE II (EC 1.9.3.1).
GN COX1/2.

```

```

OS Acanthamoeba castellanii (Amoeba).
OG Mitochondrion.
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEFF (ATCC 30010);
RX MEDLINE; 95147275.
RA BURGER G., PLANTE I., LONERGAN K.M., GRAY M.W.;
RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba
RT castellanii: complete sequence, gene content and genome
RT organization.";
RL J. Mol. Biol. 245:522-537(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NEFF (ATCC 30010);
RA BURGER G.;
RT Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O + 4
CC FERRICYTOCHROME C.
CC -!- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME
CC COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED
CC VIA HEME A AND CU(A) TO THE BINUCLEAR CENTER FORMED BY HEME A3
CC AND CU(B). SUBUNIT II BINDS CU(A) AND CYTOCHROME C.
CC -!- COFACTOR: COPPER A AND HEME GROUP.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. (MITOCHONDRIAL).
CC -!- SIMILARITY: TO OTHER MITOCHONDRIAL OR BACTERIAL COX2 SUBUNITS.
CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
DR EMBL; U12386; AAD11820.1; -.
DR PFAM; PF00115; COX1; 1.
DR PFAM; PF00116; COX2; 3.
DR PROSITE; PS00077; COX1; 1.
DR PROSITE; PS00078; COX2; 1.
KW Mitochondrion; Oxidoreductase; Copper; Transmembrane; Heme;
KW Respiratory chain.
FT METAL 803 803 COPPER A (PROBABLE).
FT METAL 807 807 COPPER A (PROBABLE).
FT METAL 836 836 COPPER A (PROBABLE).
FT METAL 840 840 COPPER A (PROBABLE).
FT METAL 844 844 COPPER A (PROBABLE).
FT METAL 847 847 COPPER A (PROBABLE).
SQ SEQUENCE 873 AA; 99213 MW; 6B4D0DD9 CRC32;

Query Match 74.5%; Score 38; DB 8; Length 873;
Best Local Similarity 87.5%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFKFFKFF 9
Db 530 FFKFSKFF 537

RESULT 5
P91319 ID P91319 PRELIMINARY; PRT; 169 AA.
AC P91319;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE COSMID F53E10.
GN F53E10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COFFEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

```

RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WORLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BECK C., WANSLEY P.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U88177; AAB42289.1; -;  
 SQ SEQUENCE 169 AA; 20087 MW; 6977591C CRC32;

Query Match 72.5%; Score 37; DB 5; Length 169;  
 Best Local Similarity 87.5%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 8  
 ||| |||||  
 DB 96 KFFKFFKFF 103

RESULT 6  
 O17265 PRELIMINARY; PRT; 592 AA.  
 ID O17265  
 AC O17265; (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DE R148.5 PROTEIN.  
 GN R148.5  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WORLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA LE T., KEMP K., SCHEET P.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF025467; AAB71039.1; -;  
 SQ SEQUENCE 592 AA; 66410 MW; DED7C7E2 CRC32;

Query Match 72.5%; Score 37; DB 5; Length 592;  
 Best Local Similarity 87.5%; Pred. No. 68;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 |||||  
 DB 36 FFKFFKFF 43

RESULT 7  
 O44959 PRELIMINARY; PRT; 548 AA.  
 ID O44959  
 AC O44959;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE M01B12.5 PROTEIN.  
 GN M01B12.5  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA BECKER M., GRAVES T., OZERSKY P.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF067624; AAC17564.1; -;  
 DR PFAM; PF01163; RIO1; 1.  
 DR PROSITE; PS01245; RIO1; 1.  
 SQ SEQUENCE 548 AA; 63781 MW; DA084B86 CRC32;

Query Match 72.5%; Score 37; DB 5; Length 548;  
 Best Local Similarity 66.7%; Pred. No. 63;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 ||:|:|  
 DB 264 KFYOFFVFF 272

RESULT 8  
 P72840 PRELIMINARY; PRT; 409 AA.  
 ID P72840  
 AC P72840;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE HYPOTHETICAL 46.6 KD PROTEIN  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RA TABATA S.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PCC6803;  
 RX MEDLINE; 97061201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,  
 RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,  
 RA TABATA S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL; D90901; BAA16855.1; -;  
 KW Hypothetical protein.

SQ SEQUENCE 409 AA; 46637 MW; 12A990FF CRC32;

Query Match 70.6%; Score 36; DB 2; Length 409;  
Best Local Similarity 66.7%; Pred. No. 70;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KFFKFFKFF 9  
DB 92 KLFPFKYF 100

RESULT 9

ID O92N10 PRELIMINARY; PRT; 414 AA.  
AC O92N10;  
DT 01-MAY-1999 (TREMREL. 10, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE EPRH.  
GN EPRH.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Staphylococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RN450;  
RA SUGAI M., FUJIWARA T., KOMATSUZAWA H., SUGINAKA H.;  
RT "Identification and molecular characterization of a gene homologous  
RT to epr (endopeptidase resistance gene) in Staphylococcus aureus.";  
RL Gene 224:67-75(1998).  
DR EMBL: AB015195; BAA33857.1; -.  
SQ SEQUENCE 414 AA; 49116 MW; E828C3D0 CRC32;

Query Match 70.6%; Score 36; DB 2; Length 414;  
Best Local Similarity 55.6%; Pred. No. 71;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KFFKFFKFF 9  
DB 61 RIFKFKYF 69

RESULT 10

ID O59824 PRELIMINARY; PRT; 709 AA.  
AC O59824;  
DT 01-AUG-1998 (TREMREL. 07, Created)  
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE PUTATIVE METALLOPEPTIDASE.  
GN SPCC965.04C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA WOOD V., RAJANDREAM M.A., BARRELL B.G., RIEGER M;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL023590; CAA19064.1; -.  
DR PFAM: PF00004; AAA: 1.  
DR PFAM: PF01434; Peptidase\_M41; 1.  
SQ SEQUENCE 709 AA; 78218 MW; 39AFBC12 CRC32;

Query Match 70.6%; Score 36; DB 3; Length 709;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KFFKFFKFF 8

DB 213 KFFRPFKF 220

RESULT 11

ID O01449 PRELIMINARY; PRT; 430 AA.  
AC O01449;  
DT 01-JUL-1997 (TREMREL. 04, Created)  
DT 01-JUL-1997 (TREMREL. 04, Last sequence update)  
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)  
DE SIMILAR TO C4-TYPE ZINC FINGER PROTEIN.  
GN C03G6.12.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA MURRAY J., WOHLDMANN P.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U97008; AAB52312.1; -.  
DR PFAM: PF00104; hormone\_rec; 1.  
DR PFAM: PF00105; zf-C4; 1.  
SQ SEQUENCE 430 AA; 50270 MW; CDB0699A CRC32;

Query Match 70.6%; Score 36; DB 5; Length 430;  
Best Local Similarity 66.7%; Pred. No. 73;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KFFKFFKFF 9  
DB 254 EYFKTFKF 262

RESULT 12

ID O18207 PRELIMINARY; PRT; 619 AA.  
AC O18207;  
DT 01-JAN-1998 (TREMREL. 05, Created)  
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE Y48E1C.3 PROTEIN.  
GN Y48E1C.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MCMURRAY A.;

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN K., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z93394; CAB07701.1; --  
 DR PFAM; PF00566; TBC; 1.  
 SQ SEQUENCE 619 AA; 72914 MW; 5A4A7D9C CRC32;

Query Match 70.6%; Score 36; DB 5; Length 619;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 ||:||| ||  
 DB 426 FFEFFAFF 433

RESULT 13

ID O96276 PRELIMINARY; PRT; 1014 AA.  
 AC O96276;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE PROTEIN WITH DNAJ DOMAIN (RESA-LIKE).  
 GN PFB0920W.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTER M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
 RT falciparum";  
 RL Science 282:1126-1132(1998).  
 DR EMBL; AE001425; AAC17973.1; --  
 DR PROSITE; PS00636; DNAJ\_1; 1.  
 SQ SEQUENCE 1014 AA; 119561 MW; E9E40338 CRC32;

Query Match 70.6%; Score 36; DB 5; Length 1014;  
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 9  
 |:||| ||:|  
 DB 2 KYFKKFF 10

RESULT 14

ID O97037 PRELIMINARY; PRT; 738 AA.  
 AC O97037;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

DE PLC-DELTAH.  
 OS Hydra magnipapillata (Hydra).  
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;  
 RC Hydridae; Hydra.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-105;  
 RA KOYANAGI M., ONO K., SUGA H., IWABE N., MIYATA T.;  
 RT "Phospholipase C cDNAs from sponge and hydra: Antiquity of genes  
 RT involved in the inositol phospholipid signaling pathway";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB017513; BAA76278.1; --  
 SQ SEQUENCE 738 AA; 84469 MW; DC8BCB99 CRC32;

Query Match 70.6%; Score 36; DB 5; Length 738;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KFFKFFKFF 8  
 :||||| |  
 DB 206 EFKFFSF 213

RESULT 15

ID O81438 PRELIMINARY; PRT; 411 AA.  
 AC O81438;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE T24H24.4 PROTEIN.  
 GN T24H24.4.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA WASHU;  
 RT "The A. thaliana Genome Sequencing Project";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA COURTNEY L., STONER K., LANGSTON Y., MEAD K.;  
 RT "The sequence of A. thaliana T24H24";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA WATERSTON R.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF075598; AAC28202.1; --  
 SQ SEQUENCE 411 AA; 47735 MW; 8B14E81A CRC32;

Query Match 70.6%; Score 36; DB 10; Length 411;  
 Best Local Similarity 75.0%; Pred. No. 70;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFKFFKFF 9  
 ||:||| ||  
 DB 127 FFEFFTF 134

Search completed: September 7, 1999, 22:47:51  
 Job time: 7971 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:47:57 ; Search time 41.46 Seconds  
(without alignments)  
14.844 Million cell updates/sec

Title: US-09-124-280A-41

Perfect score: 49

Sequence: 1 IKFLKFLKFL 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	75.5	201	051263	051263 borrelia bu
2	36	73.5	185	5 P92014	P92014 caenorhabdi
3	36	73.5	951	10 Q43106	Q43106 phaseolus v
4	36	73.5	951	10 P93265	P93265 mesembryant
5	36	73.5	956	10 Q42932	Q42932 nicotiana p
6	36	73.5	963	10 Q43131	Q43131 vicia faba
7	36	73.5	956	10 Q43182	Q43182 solanum tub
8	36	73.5	949	10 Q43243	Q43243 zea mays (m
9	36	73.5	956	10 Q43001	Q43001 oryza sativ
10	36	73.5	957	10 Q43002	Q43002 oryza sativ
11	35	71.4	1289	5 Q17174	Q17174 boophilus m
12	35	71.4	72	5 Q23632	Q23632 caenorhabdi
13	35	71.4	1278	5 O96228	O96228 plasmodium
14	34	69.4	564	2 O67184	O67184 aquifex aeo
15	34	69.4	1119	5 O77374	O77374 plasmodium
16	34	69.4	1687	11 Q61204	Q61204 mus musculu
17	33	67.3	644	2 O67389	O67389 aquifex aeo
18	33	67.3	141	2 O67910	O67910 aquifex aeo
19	33	67.3	53	2 O68585	O68585 streptococ
20	33	67.3	45	4 Q14460	Q14460 homo sapien
21	33	67.3	93	5 O01587	O01587 caenorhabdi
22	33	67.3	686	10 Q40940	Q40940 chlorarachn
23	33	67.3	4572	10 Q40712	Q40712 oryza sativ
24	33	67.3	72	12 Q88417	Q88417 spiroplasma
25	32	65.3	129	1 O29640	O29640 archaeoglob
26	32	65.3	260	1 O27551	O27551 methanobact
27	32	65.3	132	1 O58773	O58773 pyrococcus
28	32	65.3	130	2 O67445	O67445 aquifex aeo
29	32	65.3	140	2 O32317	O32317 bacillus th

#### ALIGNMENTS

##### RESULT 1

051263 PRELIMINARY; PRT; 201 AA.  
ID OS1263  
AC OS1263  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE CONSERVED HYPOTHETICAL PROTEIN.  
GN BB0247.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE; 98065943.  
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
RA PETERSON J., KERLAVAGE A.R., OUACKENBUSH J., SALZBERG S., HANSON M.,  
RA VAN VUUT J., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
RA SMITH H.O., VENTER J.C.;  
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."  
RL Nature 390:580-586(1997).  
DR EMBL; AE001135; AAB91499.1; -.  
DR TIGR; BB0247; -.  
SQ SEQUENCE 201 AA; 23138 MW; 1F87F075 CRC32;

Query Match 75.5%; Score 37; DB 2; Length 201;  
Best Local Similarity 80.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

##### QY 1 IKFLKFLKFL 10

Db 185 IAFKFKFL 194

##### RESULT 2

P92014 PRELIMINARY; PRT; 185 AA.  
ID P92014  
AC P92014  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
DE T02B5.2 PROTEIN.  
GN T02B5.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
OC Rhabditina; Rhabditidae; Peloderinae; Caenorhabditis.

RN SEQUENCE FROM N.A.  
RP MCURRAY A.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL Nature 368:32-38(1994).  
DR EMBL: Z81112; CAB03274.1; -;  
SQ SEQUENCE 185 AA; 21299 MW; A87A13B5 CRC32;

Query Match 73.5%; Score 36; DB 5; Length 185;  
Best Local Similarity 70.0%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
| | | | |  
Db 169 INFLNKLFL 178

RESULT 3  
Q43106 PRELIMINARY; PRT; 951 AA.  
AC Q43106;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
GN BHA-1.  
OS Phaseolus vulgaris (Kidney bean) (French bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
OC Phaseolus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA CAMPOS F., PEREZ-CASTINEIRA J.R., VILLALBA J.M., SERRANO R.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: X85804; CAA59799.1; -;  
DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
DR MENDEL; 13288; PHAVU:494;1.  
DR PFAM: PF00122; E1-E2\_ATPase; 1.  
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
FT MOD\_RES 331 331 PHOSPHORYLATION (PROBABLE).  
SQ SEQUENCE 951 AA; 104442 MW; 3B569E8C CRC32;

Query Match 73.5%; Score 36; DB 10; Length 951;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
| | | | |  
Db 59 KFLKFLGFM 67

RESULT 4  
P93265 PRELIMINARY; PRT; 953 AA.  
ID P93265  
AC P93265;

DT 01-MAY-1997 (TRENBLrel. 03, Created)  
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)  
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
GN PMA.  
OS Mesembryanthemum crystallinum (Common ice plant).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Caryophyllales; Caryophyllaceae; Aizoaceae;  
OC Mesembryanthemum.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MICHALOWSKI C.B., QUIGLEY-LANDREAU F., BOHNERT H.J.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U84891; ABA41898.1; -;  
DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
DR MENDEL; 13177; MESCR:494;1.  
DR PFAM: PF00122; E1-E2\_ATPase; 1.  
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
FT MOD\_RES 335 335 PHOSPHORYLATION (PROBABLE).  
SQ SEQUENCE 953 AA; 104924 MW; B11E9AA4 CRC32;

Query Match 73.5%; Score 36; DB 10; Length 953;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
| | | | |  
Db 63 KFLKFLGFM 71

RESULT 5  
Q42932 PRELIMINARY; PRT; 956 AA.  
AC Q42932;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
DE H(+)-TRANSPORTING ATPASE.  
OS Nicotiana glauca (Leadwort-leaved tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
OC Nicotiana.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89334860.  
RA BOUTRY M., MICHELET B., GOFFEAU A.;  
RL "Molecular cloning of a family of plant genes encoding a protein  
RT homologous to plasma membrane H<sup>+</sup>-translocating ATPases.";  
RL Biochem. Biophys. Res. Commun. 162:567-574(1989).  
DR EMBL: M27886; AAA34052.1; -;  
DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
DR MENDEL; 16027; NICPL:494;mn16027.  
DR PFAM: PF00122; E1-E2\_ATPase; 1.  
KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
SQ SEQUENCE 956 AA; 105047 MW; FF2D51BE CRC32;

Query Match 73.5%; Score 36; DB 10; Length 956;  
Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
| | | | |  
Db 61 KFLKFLGFM 69

RESULT 6  
Q43131 PRELIMINARY; PRT; 963 AA.  
ID Q43131

AC Q43131;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H(+)-ATPASE.  
 OS Vicia faba (Broad bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eumyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;  
 OC Vicia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96012934.  
 RA NAKAJIMA N., SAJI H., AONO M., KONDO N.;  
 RT "Isolation of cDNA for a plasma membrane H(+)-ATPase from guard cells  
 of Vicia faba L.";  
 RL Plant Cell Physiol. 36:919-924(1995).  
 DR EMBL; S79323; AAB35314.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13431; VICfa;494;3.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.  
 DR KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 338 338 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 963 AA; 105965 MW; 2C2C9793 CRC32;

Query Match 73.5%; Score 36; DB 10; Length 963;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KFLKFLKFL 10  
 ||||| 1;  
 DB 66 KFLKFLGFM 74

RESULT 7

Q43182 ID Q43182 PRELIMINARY; PRT; 956 AA.  
 AC Q43182;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE (EC 3.6.1.35).  
 GN PHA1.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eumyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
 OC Solanum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=LEAF;  
 RX MEDLINE; 95093037.  
 RA HARMS K., WOHRER R.V., SCHULZ B., FROMMER W.B.;  
 RT "Isolation and characterization of P-type H(+)-ATPase genes from  
 potato.";  
 RL Plant Mol. Biol. 26:979-988(1994).  
 DR EMBL; X76536; CAA54046.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 530; SOLtu;494;1.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.  
 DR KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105125 MW; 80ADD145 CRC32;

Query Match 73.5%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KFLKFLKFL 10  
 ||||| 1;  
 DB 61 KFLKFLGFM 69

RESULT 8  
 Q43243 ID Q43243 PRELIMINARY; PRT; 949 AA.  
 AC Q43243;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
 DE H(+)-TRANSPORTING ATPASE.  
 GN MH1.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eumyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Zea.  
 RN [1]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RX STRAIN=D3L;  
 RX MEDLINE; 95003707.  
 RA JIN Y., BENNETZEN J.L.;  
 RT "Integration and nonrandom mutation of a plasma membrane proton  
 ATPase gene fragment within the Bsl retroelement of maize.";  
 RL Plant Cell 6:1177-1186(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=D3L;  
 RC JIN Y.;  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U09989; AAB60276.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 533; ZEama;494;2.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.  
 DR KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 336 336 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 949 AA; 103900 MW; DDIAA0CF CRC32;

Query Match 73.5%; Score 36; DB 10; Length 949;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KFLKFLKFL 10  
 ||||| 1;  
 DB 64 KFLKFLGFM 72

RESULT 9

Q43001 ID Q43001 PRELIMINARY; PRT; 956 AA.  
 AC Q43001;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE H-ATPASE (EC 3.6.1.3) (ADENOSINETRIPHOSPHATASE)  
 DE (ADENYLPHOSPHATASE) (ATP MONOPHOSPHATASE) (TRIPHOSPHATASE).  
 GN OSAL.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC eumyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=NIPPONBARE K-1;  
 RA WADA M., TAKANO M., KASAMO K.;  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.  
 DR EMBL; D10207; BAA01058.1; -;  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR MENDEL; 13241; Oryza;494;1.  
 DR PFAM; PF00122; El-E2\_ATPase; 1.  
 DR KW Hydrolase; Transmembrane; Phosphorylation; ATP-binding.  
 FT MOD\_RES 333 333 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 956 AA; 105079 MW; 1A9E387B CRC32;

Query Match 73.5%; Score 36; DB 10; Length 956;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 :|||||:  
 Db 61 KFLKFLGFM 69

RESULT 10  
 Q43002 PRELIMINARY; PRT; 957 AA.  
 AC Q43002:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE PLASMA MEMBRANE H+-ATPASE.  
 GN OSA2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euryliophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 OC Poaceae; Oryza.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RX MEDLINE; 95211383.  
 RA OKURA T., WADA M., SAKAKIBARA Y., JEONG K.H., MARUTA I., KAWAMURA Y.,  
 RA KASANO K.;  
 RT "Identification and characterization of a family of genes for the  
 RT plasma membrane H(+)-ATPase of Oryza sativa L.";  
 RL Plant Cell Physiol. 35:1251-1256(1994).  
 DR EMBL; D31843; BAA06629.1; -.  
 DR MENDEL; 13242; Orysa;494;2.  
 DR PFAM; PF00122; E1-E2 ATPase; 1.  
 SQ SEQUENCE 957 AA; 105607 MW; 6673B764 CRC32;

Query Match 73.5%; Score 36; DB 10; Length 957;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
 :|||||:  
 Db 61 KFLKFLGFM 69

RESULT 11  
 Q17174 PRELIMINARY; PRT; 1289 AA.  
 AC Q17174:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE GP80 (FRAGMENT).  
 OS Boophilus microplus (Cattle tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
 RN [1]  
 RP SEQUENCE FROM N.A..  
 RC STRAIN=YEERONGPILLY;  
 RA IRVING D., SMITH D.R.J., SHARP P., RIDING G., BRISCOE S.,  
 RA TELLAM R.L.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U49934; AAA92143.1; -.  
 FT NON-TER 1  
 SQ SEQUENCE 1289 AA; 147185 MW; A82B0E07 CRC32;

Query Match 71.4%; Score 35; DB 5; Length 1289;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKF 9  
 :||||:  
 Db 889 LKFLSYLKF 897

RESULT 12  
 Q23632 PRELIMINARY; PRT; 72 AA.  
 AC Q23632:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE COSMID ZK84.  
 GN ZK84.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditinae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA KIRSTEN J.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U23181; AAC48206.1; -.  
 SQ SEQUENCE 72 AA; 8643 MW; AD8F8726 CRC32;

Query Match 71.4%; Score 35; DB 5; Length 72;  
 Best Local Similarity 77.8%; Pred. No. 18;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKF 9  
 :||||:  
 Db 45 LKFLFLKLF 53

RESULT 13  
 Q96228 PRELIMINARY; PRT; 1278 AA.  
 AC Q96228:  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE PREDICTED SECRETED PROTEIN.  
 GN PFB0675W.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTSEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT Chromosome 2 sequence of the human malaria parasite Plasmodium  
 RT falciparum.";

RL Science 282:1126-1132(1998).  
 DR EMBL; AE001410; AAC71924.1; -.  
 SQ SEQUENCE 1278 AA; 149634 MW; 9A4F5FED CRC32;

Query Match 71.4%; Score 35; DB 5; Length 1278;  
 Best Local Similarity 70.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IKFLKFLKFL 10  
 ||| ||| |  
 Db 687 IKCKELLFI 696

## RESULT 14

O67184 PRELIMINARY; PRT; 564 AA.  
 AC O67184;  
 DT 01-AUG-1998 (TReMBLrel. 07, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE DE ABC TRANSPORTER (HLYB SUBFAMILY).  
 GN ABCT5.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus."  
 RT Nature 392:353-358(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE000723; AAC07149.1; -.  
 DR PFAM; PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 564 AA; 63100 MW; 9716784B CRC32;

Query Match 69.4%; Score 34; DB 2; Length 564;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KFLKFLKF 9  
 :|||:|  
 Db 310 EFLKFLRE 317

## RESULT 15

O77374 PRELIMINARY; PRT; 1119 AA.  
 AC O77374;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MAL3P6.21 PROTEIN.  
 GN MAL3P6.21.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-3D7;  
 RA MURPHY L., LAWSON D., BARRELL B.;

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z98551; CAB11130.1; -.  
 SQ SEQUENCE 1119 AA; 132982 MW; 33F6CED5 CRC32;

Query Match 69.4%; Score 34; DB 5; Length 1119;  
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IKFLKFLKFL 10  
 ||| :|||  
 Db 932 IKFINFEFL 941

Search completed: September 7, 1999, 23:17:30  
 Job time: 1773 sec

**THIS PAGE BLANK (USPTO)**

P41308 mycoplasma  
P20649 arabidopsis

## ALIGNMENTS

RESULT	1
PMAL_LYCES	
ID	PMAL_LYCES
AC	STANDARD;
P22180;	PRT;
DT	956 AA.
DT	01-AUG-1991 (REL. 19, CREATED)
DT	01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE	DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP)

## References

OS LYCOPERSICON ESCULENTUM (TOMATO).  
OC EUKARYOTA: VIRIDIPANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC ASTRIDAE: SOLANACE: SOLANALES: SOLANACEAE: SOLANUM

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

•

-1- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE ACTIVE TRANSPORT OF NUTRIENTS BY H<sup>+</sup>-SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKALINIZATION MAY MEDIATE GROWTH RESPONSES.  
 -1- CATALYTIC ACTIVITY: ATP + H<sub>2</sub>O = ADP + ORTHOPHOSPHATE.  
 -1- SUBUNIT: POSSIBLY EXISTS AS AN HOMODIMER OR AN HOMOTRIMER.  
 -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 -1- AS MANY AS 6 TO 8 CLOSELY RELATED GENES MAY ENCODE OTHER ISOFORMS OF PLASMA MEMBRANE ATPASE IN TOMATO, LIKE THE LHA2 GENE PRODUCT WHICH IS 96% IDENTICAL TO THE LHA1 GENE PRODUCT.  
 -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES).

-----  
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; M60166; G170464; -  
 PIR; A45506; A45506.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 DR PFAM; PF00122; E1-E2\_ATPase; 1.  
 KW HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;  
 KW ATP-BINDING; MULTIGENE FAMILY.  
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 65 85 1 (POTENTIAL).  
 FT DOMAIN 86 96 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 97 116 2 (POTENTIAL).  
 FT DOMAIN 117 245 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 246 267 3 (POTENTIAL).  
 FT DOMAIN 268 279 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 280 302 4 (POTENTIAL).  
 FT DOMAIN 303 677 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 678 696 5 (POTENTIAL).  
 FT DOMAIN 697 712 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 713 736 6 (POTENTIAL).  
 FT DOMAIN 737 794 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 795 815 7 (POTENTIAL).  
 FT DOMAIN 816 830 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 831 847 8 (POTENTIAL).  
 FT DOMAIN 848 956 EXTRACELLULAR (POTENTIAL).  
 FT MOD\_RES 333 333 PHOSPHORYLATION (BY SIMILARITY).  
 FT BINDING 427 427 ATP (BY SIMILARITY).

SQ SEQUENCE 956 AA; 105103 MW; E798A2DA CRC32;

Query Match 73.5%; Score 36; DB 1; Length 956;  
Best Local Similarity 77.8%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
|||||  
Db 61 KFLKFLGFM 69

## RESULT 2

PSP2\_YEAST STANDARD; PRT; 578 AA.  
AC PS0109; Q071170;  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PSP2 PROTEIN (MITOCHONDRIAL REGULATOR OF SPLICING 15).  
GN PSP2 OR MRS15 OR YML017W OR YN9571.O1.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA FORMOSA T.D., NITTIS T.;  
RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSY747;  
RA TEPLY R.;  
RL THESIS (1995), VIENNA BIOCENTRE, AUSTRIA.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA GENTLES S., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: DNA POLYMERASE ALPHA MUTATION SUPPRESSOR.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U33116; G995753; .  
DR EMBL; U29398; G1808701; .  
DR EMBL; Z49810; G854473; .  
DR SGD; L0002882; PSP2.  
FT DOMAIN 393 402 POLY-ASN.  
FT DOMAIN 438 453 POLY-ASN.  
FT DOMAIN 457 460 POLY-ASN.  
SQ SEQUENCE 578 AA; 64020 MW; 7DA753A1 CRC32;

Query Match 69.4%; Score 34; DB 1; Length 578;  
Best Local Similarity 75.0%; Pred. No. 31;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKF 9  
|||  
Db 117 KFTKFKFK 124

## RESULT 3

ALR2\_BACSU STANDARD; PRT; 394 AA.  
ID ALR2\_BACSU  
AC P94494;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE PUTATIVE ALANINE RACEMASE (EC 5.1.1.1).  
GN YNCD.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BORCHERT S., KLEIN C., PIKSA B., HAMMELMANN M., ENTIAN K.D.;  
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- CATALYTIC ACTIVITY: L-ALANINE = D-ALANINE.  
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -!- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U66480; G1750128; .  
DR EMBL; Z99113; E1183423; .  
DR SUBTILIST; BG12267; YNCD.  
DR PROSITE; PS00395; ALANINE RACEMASE; 1.  
DR PRAM; PF00842; Ala\_racemase; 1.  
KW HYPOTHETICAL PROTEIN; ISOMERASE; PYRIDOXAL PHOSPHATE.  
FT BINDING 39 39 PYRIDOXAL PHOSPHATE.  
SQ SEQUENCE 394 AA; 43648 MW; B5C38A51 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 394;  
Best Local Similarity 66.7%; Pred. No. 32;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10  
|||  
Db 187 KFISFLSFL 195

## RESULT 4

RT07\_ACACA STANDARD; PRT; 337 AA.  
AC P45756; 1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL RIBOSOMAL PROTEIN S7.  
GN RPS7.  
OS ACANTHAMOEBA CASTELLANII (AMOEBA).  
OC MITOCHONDRION.  
OC EUKARYOTA; ACANTHAMOEBIIDAE; ACANTHAMOEBA.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U12386; G562050; .  
DR PROSITE; PS00052; RIBOSOMAL\_S7; FALSE\_NEG.



```

DR PFAM; PF00177; S7; 1.
KW RIBOSOMAL PROTEIN; MITOCHONDRION.
SQ SEQUENCE 337 AA; 41612 MW; 9817F1E4 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 337;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLK 8
   |::|::|
Db 327 IRLRLFLK 334

RESULT 5
VGNB_CPSMV STANDARD; PRT; 1858 AA.
AC P36312; 1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN B [CONTAINS: PROTEASE COFACTOR; MEMBRANE BINDING
DE PROTEIN: VPG; PROTEASE (EC 3.4.22.-); RNA POLYMERASE (EC 2.7.7.48)].
OS CORPEA SEVERE MOSAIC VIRUS (STRAIN DG) (CPSMV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; COMOVIRIDAE;
OC COMOVIRUS.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE; 93079863.
RA CHEN X., BRUENING G.;
RT "Cloned DNA copies of cowpea severe mosaic virus genomic RNAs:
RT infectious transcripts and complete nucleotide sequence of RNA 1.";
RL VIROLOGY 191:607-618(1992).
CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M83830; -. NOT_ANNOTATED_CDS.
DR PIR; A44214; A44214.
DR PFAM; PF00680; RNA_dep_RNA_pol; 1.
DR PFAM; PF00910; RNA_helicase; 1.
KW POLYPROTEIN; TRANSMEMBRANE; HYDROLASE; PROTEASE; TRANSFERASE;
KW RNA-DIRECTED RNA POLYMERASE; ATP-BINDING.
FT CHAIN 1 313 PROTEASE COFACTOR (POTENTIAL).
   |::|::|
FT CHAIN 314 908 MEMBRANE-BINDING PROTEIN (POTENTIAL).
   |::|::|
FT CHAIN 909 936 VPG PROTEIN (POTENTIAL).
   |::|::|
FT CHAIN 937 1146 PROTEASE (POTENTIAL).
FT CHAIN 1147 1858 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
FT TRANSMEM 891 907 ATP (POTENTIAL).
FT NP_BIND 482 489 THIOL PROTEASE (POTENTIAL).
FT ACT_SITE 976 976 THIOL PROTEASE (POTENTIAL).
FT ACT_SITE 1013 1013 THIOL PROTEASE (POTENTIAL).
FT ACT_SITE 1104 1104 THIOL PROTEASE (POTENTIAL).
SQ SEQUENCE 1858 AA; 208861 MW; 38B826AB CRC32;

Query Match 67.3%; Score 33; DB 1; Length 1858;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFLK 10
   |::|::|
Db 1379 KFLKFLKFLK 1387

RESULT 6
YGNB_CPSMV STANDARD; PRT; 1122 AA.
AC P53280;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 128.8 KD PROTEIN IN PAS2-PRS5 INTERGENIC REGION.
GN YGR134W.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
SEQUENCE FROM N.A.
RA VAN DYCK L., SKALA J., DE WERGIFOSSE P., PURNELLE B., TALLA E.,
RA NAWROCKI A., DEL BINO S., GOFFEAU A.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; Z72919; E243526; -.
DR KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 1122 AA; 128779 MW; 307CEC10 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 1122;
Best Local Similarity 70.0%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKFLK 10
   |::|::|
Db 125 ILLKFLRLFL 134

RESULT 7
YJ20_YEAST STANDARD; PRT; 110 AA.
AC P47092;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 12.3 KD PROTEIN IN ESS1-MER2 INTERGENIC REGION.
GN YJR020W OR J1458 OR YJR83.15.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RA DE HAAN M., SMITS P.H.M., GRIVELL L.A.;
RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X87611; G854593; -.
DR EMBL; Z49519; G1015656; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 110 AA; 12281 MW; 05F9AAC2 CRC32;

Query Match 67.3%; Score 33; DB 1; Length 110;
Best Local Similarity 66.7%; Pred. No. 9.5;

```

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IKFLKFLKF 9  
:|||||:  
Db 1 MKFLPFLRF 9

RESULT 8  
Y037\_METJA STANDARD; PRT; 238 AA.  
AC Q60344;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN MJ0037.  
GN MJ0037.  
OS METHANOCOCCUS JANNASCHII.  
OC ARCHAEA; EURVARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
OC METHANOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE; 96337999.  
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,  
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,  
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.,  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL SCIENCE 273:1058-1073(1996).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U57462; G1498797;  
DR TIGR; MJ0037;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 238 AA; 27684 MW; C5CBD35C CRC32;  
-----  
Query Match 66.3%; Score 32.5; DB 1; Length 238;  
Best Local Similarity 72.7%; Pred. No. 25;  
Matches 8; Conservative 2; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 IKFLK-FLKFL 10  
|||||:|:  
Db 85 IKFLKEFIEFL 95

RESULT 9  
DIP2\_YEAST STANDARD; PRT; 943 AA.  
ID DIP2\_YEAST  
AC Q12220; Q05386;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DOM34 INTERACTING PROTEIN 2.  
GN DIP2 OR YLR129W OR L9233.1 OR L3116.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
-----  
Query Match 65.3%; Score 32; DB 1; Length 943;  
Best Local Similarity 85.7%; Pred. No. 1.e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 LKFLKFL 10  
|||||:  
Db 830 LKFLKFI 836

RESULT 10  
NU5M\_ARBLI STANDARD; PRT; 190 AA.  
ID NU5M\_ARBLI  
AC Q33753; Q33755;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3) (FRAGMENT).  
GN ND5.  
OS ARBACIA LIXULA (BLACK URCHIN).  
OG MITOCHONDRION.  
OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOIDEA;

RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,  
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,  
RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,  
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,  
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,  
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,  
RA WILSON R., WATERSTON R.,  
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA DELIUS H.,  
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 1-458 FROM N.A.  
RC STRAIN-FY23 / RD005;  
RA VERHASSELT P., VOET M., VOLCKAERT G.,  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -I- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).  
CC -I- SIMILARITY: TO S.POMBE SPBC3D5.12.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U53877; G1256869;  
DR EMBL; U53881; G2580465;  
DR EMBL; X91258; G995691;  
DR EMBL; Z73301; E245568;  
DR EMBL; Z73302; E322129;  
DR EMBL; X89514; E198829;  
DR SGD; L0003088; DIP2.  
DR PROSITE; PS00678; WD\_REPEATS; 4.  
DR PRAM; PF00400; G-beta; 8.  
KW REPEAT; WD REPEAT.  
FT REPEAT 77 107  
FT REPEAT 119 149  
FT REPEAT 161 190  
FT REPEAT 202 230  
FT REPEAT 389 418  
FT REPEAT 428 458  
FT REPEAT 471 501  
FT REPEAT 571 601  
FT REPEAT 613 643  
FT REPEAT 655 685  
SQ SEQUENCE 943 AA; 106342 MW; 7B348EA8 CRC32;  
-----  
Query Match 65.3%; Score 32; DB 1; Length 943;  
Best Local Similarity 85.7%; Pred. No. 1.e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 LKFLKFL 10  
|||||:  
Db 830 LKFLKFI 836

RESULT 10  
NU5M\_ARBLI STANDARD; PRT; 190 AA.  
ID NU5M\_ARBLI  
AC Q33753; Q33755;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3) (FRAGMENT).  
GN ND5.  
OS ARBACIA LIXULA (BLACK URCHIN).  
OG MITOCHONDRION.  
OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOIDEA;

OC EUECHINOIDEA; ECHINACEA; ARBACOIDA; ARBACIIDAE; ARBACIA.  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92017217.  
 RA DE GIORGI C., LANAVE C., MUSCI M.D., SACCONI C.;  
 RT "Mitochondrial DNA in the sea urchin Arbacia lixula: evolutionary  
 RT inferences from nucleotide sequence analysis."  
 RL MOL. BIOL. EVOL. 8:515-529(1991).  
 RN [2]  
 RP SEQUENCE OF 1-80 FROM N.A.  
 RX MEDLINE; 91365253.  
 RA DE GIORGI C., DE LUCA F., SACCONI C.;  
 RT "Mitochondrial DNA in the sea urchin Arbacia lixula: nucleotide  
 RT sequence differences between two polymorphic molecules indicate  
 RT asymmetry of mutations."  
 RL GENE 103:249-252(1991).  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M74839; G1280184; -.  
 DR EMBL; M79454; E42825; -.  
 DR PFAM; PF00662; oxidored\_g1.n: 1.  
 KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION.  
 FT VARIANT 14 14 L -> S.  
 FT NON\_TER 190 190  
 SQ SEQUENCE 190 AA; 21157 MW; 74811B8D CRC32;  
 Query Match 65.38; Score 32; DB 1; Length 190;  
 Best Local Similarity 70.08; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 IKFLKFLKFL 10  
 I | | | | I |  
 DB 68 INILKFLAFL 77  
 RESULT 11  
 NU6M\_CVACA STANDARD; PRT; 201 AA.  
 ID NU6M\_CVACA  
 AC P48925;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 6 (EC 1.6.5.3).  
 GN ND6 OR NAD6.  
 OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).  
 OC MITOCHONDRION.  
 CC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDIAEAE;  
 CC CYANIDIUM.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 RP STRAIN-RK-1;  
 RA VIEHMANN S.;  
 RL THESIS (1995), JUSTUS LIEBIG UNIVERSITAET / FRANKFURT, GERMANY.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z48930; G758173; -.

DR PFAM; PF00499; oxidored\_q3; 1.  
 DR MENDEL; 7852; CYACA; nad6; 1.  
 KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION.  
 SQ SEQUENCE 201 AA; 23280 MW; 4AA62495 CRC32;  
 Query Match 65.38; Score 32; DB 1; Length 201;  
 Best Local Similarity 75.08; Pred. No. 26;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 IKFLKFLK 8  
 I | | | | I |  
 DB 194 IRFFKFLK 201  
 RESULT 12  
 PEXB\_YEAST STANDARD; PRT; 235 AA.  
 ID PEXB\_YEAST  
 AC Q12462;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE PEROXISOMAL MEMBRANE PROTEIN PMP27 (PEROXIN-11).  
 GN PEX11 OR PMP27 OR PMP24 OR YOL147C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 CC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 CC [1]  
 CC SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.  
 RP STRAIN-S288C;  
 RX MEDLINE; 95164555.  
 RA ERDMANN R., BLOBEL G.;  
 RT "Giant peroxisomes in oleic acid-induced Saccharomyces cerevisiae  
 RT lacking the peroxisomal membrane protein Pmp27p."  
 RL J. CELL BIOL. 128:509-523(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-29 AND 191-206.  
 RC STRAIN-GRF88;  
 RX MEDLINE; 95238534.  
 RA MARSHALL P.A., KRIMKEVICH Y.I., LARK R.H., DYER J.M., VEENHUIS M.,  
 RA GOODMAN J.M.;  
 RT "Pmp27 promotes peroxisomal proliferation."  
 RL J. CELL BIOL. 129:345-355(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / FY1679;  
 RX MEDLINE; 96132030.  
 RA CASAMAYOR A., ALDEA M., CASAS C., HERRERO E., GAMO F.J.,  
 RA LAFUENTE M.J., GANCEDO C., ARINO J.;  
 RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast  
 RT chromosome XV containing seven new open reading frames."  
 RL YEAST 11:1281-1288(1995).  
 CC -1- FUNCTION: INVOLVED IN PEROXISOMAL PROLIFERATION. COULD PARTICIPATE  
 CC IN PEROXISOMAL ELONGATION OR FISSION. MAY BE INVOLVED IN PARCELING  
 CC OF PEROXISOMES INTO REGULAR QUANTA.  
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL MEMBRANE-ASSOCIATED.  
 CC -1- SIMILARITY: STRONG, TO C.BOIDINII PEX11.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z48239; G886949; -.  
 DR EMBL; Z74889; E251925; -.  
 DR EMBL; X81465; G683701; -.  
 DR EMBL; Z45846; G791132; -.  
 DR SGD; L0002650; PMP27.  
 KW PEROXISOME; MEMBRANE.  
 FT INIT\_MET 0

SQ SEQUENCE 235 AA; 26744 MW; 228BDA38 CRC32;

Query Match 65.3%; Score 32; DB 1; Length 235;  
Best Local Similarity 85.7%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLK 8  
11111111  
Db 59 KFLK 65

RESULT 13

SUR\_HUMAN STANDARD; PRT; 1580 AA.

AC Q09428;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE SULFONYLUREA RECEPTOR.

GN SUR  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

RA GONZALEZ G., AGUILAR-BRYAN L., BRYAN J.;  
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

RP SEQUENCE FROM N.A.

RC TISSUE-PANCREATIC ISLETS;

RX MEDLINE; 95232533.

RA THOMAS P.M., COTE G.J., WOHLK N., HADDAD B., MATHEW P.M., RABL W.,  
AGUILAR-BRYAN L., GAGEL R.F., BRYAN J.;  
RT "Mutations in the sulfonylurea receptor gene in familial persistent  
hyperinsulinemic hypoglycemia of infancy.";  
RL SCIENCE 268:426-429(1995).

RN [1]

RP VARIANT PPHI VAL-717.

RX MEDLINE; 96354544.

RA THOMAS P.M., WOHLK N., HUANG E., KUNLE U., RABL W., GAGEL R.F.,  
COTE G.J.;  
RT "Inactivation of the first nucleotide-binding fold of the  
sulfonylurea receptor, and familial persistent hyperinsulinemic  
hypoglycemia of infancy.";  
RL AM. J. HUM. GENET. 59:510-518(1996).

CC -!- FUNCTION: PUTATIVE SUBUNIT OF THE BETA-CELL ATP-SENSITIVE  
POTASSIUM CHANNEL (KATP). REGULATOR OF ATP-SENSITIVE K+ CHANNELS  
AND INSULIN RELEASE.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -!- DISEASE: MUTATIONS IN THE SUR PROTEIN ARE THE CAUSE OF FAMILIAL  
PERSISTENT HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY (PHHI), AN  
AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED BY UNREGULATED INSULIN  
SECRETION.

CC -!- DISEASE: DEFECTS IN SUR ARE A CAUSE OF FAMILIAL PERSISTENT  
HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY (PHHI), AN AUTOSOMAL  
RECESSIVE DISORDER CHARACTERIZED BY UNREGULATED INSULIN SECRETION.  
IT CAUSES NESTROBLASTOSIS, A DIFFUSE ABNORMALITY OF THE PANCREAS  
IN WHICH THERE IS EXTENSIVE, OFTEN DISORGANIZED FORMATION OF NEW  
ISLETS.

CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
(ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; L78243; G1374919; JOINED.

CC EMBL; L78208; G1374919; JOINED.

DR EMBL; L78209; G1374919; JOINED.  
DR EMBL; L78210; G1374919; JOINED.  
DR EMBL; L78211; G1374919; JOINED.  
DR EMBL; L78212; G1374919; JOINED.  
DR EMBL; L78255; G1374919; JOINED.  
DR EMBL; L78213; G1374919; JOINED.  
DR EMBL; L78214; G1374919; JOINED.  
DR EMBL; L78215; G1374919; JOINED.  
DR EMBL; L78216; G1374919; JOINED.  
DR EMBL; L78217; G1374919; JOINED.  
DR EMBL; L78218; G1374919; JOINED.  
DR EMBL; L78219; G1374919; JOINED.  
DR EMBL; L78220; G1374919; JOINED.  
DR EMBL; L78221; G1374919; JOINED.  
DR EMBL; L78222; G1374919; JOINED.  
DR EMBL; L78223; G1374919; JOINED.  
DR EMBL; L78225; G1374919; JOINED.  
DR EMBL; L78254; G1374919; JOINED.  
DR EMBL; L78226; G1374919; JOINED.  
DR EMBL; L78227; G1374919; JOINED.  
DR EMBL; L78228; G1374919; JOINED.  
DR EMBL; L78229; G1374919; JOINED.  
DR EMBL; L78230; G1374919; JOINED.  
DR EMBL; L78231; G1374919; JOINED.  
DR EMBL; L78232; G1374919; JOINED.  
DR EMBL; L78233; G1374919; JOINED.  
DR EMBL; L78234; G1374919; JOINED.  
DR EMBL; L78235; G1374919; JOINED.  
DR EMBL; L78236; G1374919; JOINED.  
DR EMBL; L78237; G1374919; JOINED.  
DR EMBL; L78238; G1374919; JOINED.  
DR EMBL; L78239; G1374919; JOINED.  
DR EMBL; L78240; G1374919; JOINED.  
DR EMBL; L78241; G1374919; JOINED.  
DR EMBL; L78242; G1374919; JOINED.  
DR EMBL; L78243; G1374919; JOINED.  
DR EMBL; L78208; G1374920; JOINED.  
DR EMBL; L78209; G1374920; JOINED.  
DR EMBL; L78210; G1374920; JOINED.  
DR EMBL; L78211; G1374920; JOINED.  
DR EMBL; L78212; G1374920; JOINED.  
DR EMBL; L78255; G1374920; JOINED.  
DR EMBL; L78213; G1374920; JOINED.  
DR EMBL; L78214; G1374920; JOINED.  
DR EMBL; L78215; G1374920; JOINED.  
DR EMBL; L78216; G1374920; JOINED.  
DR EMBL; L78217; G1374920; JOINED.  
DR EMBL; L78218; G1374920; JOINED.  
DR EMBL; L78219; G1374920; JOINED.  
DR EMBL; L78220; G1374920; JOINED.  
DR EMBL; L78221; G1374920; JOINED.  
DR EMBL; L78222; G1374920; JOINED.  
DR EMBL; L78223; G1374920; JOINED.  
DR EMBL; L78225; G1374920; JOINED.  
DR EMBL; L78254; G1374920; JOINED.  
DR EMBL; L78226; G1374920; JOINED.  
DR EMBL; L78227; G1374920; JOINED.  
DR EMBL; L78228; G1374920; JOINED.  
DR EMBL; L78229; G1374920; JOINED.  
DR EMBL; L78230; G1374920; JOINED.  
DR EMBL; L78231; G1374920; JOINED.  
DR EMBL; L78232; G1374920; JOINED.  
DR EMBL; L78233; G1374920; JOINED.  
DR EMBL; L78234; G1374920; JOINED.  
DR EMBL; L78235; G1374920; JOINED.  
DR EMBL; L78236; G1374920; JOINED.  
DR EMBL; L78237; G1374920; JOINED.  
DR EMBL; L78238; G1374920; JOINED.  
DR EMBL; L78239; G1374920; JOINED.  
DR EMBL; L78240; G1374920; JOINED.  
DR EMBL; L78241; G1374920; JOINED.  
DR EMBL; L78242; G1374920; JOINED.  
DR EMBL; L78207; G1369844; JOINED.

DR EMBL; L40625; G784882; --  
 DR MM; 600509; --  
 DR MM; 256450; --  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 DR PFAM; PF00005; ABC\_tran; 2.  
 DR PFAM; PF00664; ABC\_membrane; 2.  
 DR HSP; P13569; INBD.  
 KW RECEPTOR; TRANSMEMBRANE; ATP-BINDING; PHOSPHORYLATION; GLYCOPROTEIN;  
 ALTERNATIVE SPLICING; DISEASE MUTATION.  
 FT INIT\_MET 0  
 FT DOMAIN 1 29  
 FT TRANSSEM 30 50  
 FT DOMAIN 51 73  
 FT TRANSSEM 74 93  
 FT DOMAIN 94 137  
 FT TRANSSEM 138 153  
 FT DOMAIN 154 167  
 FT TRANSSEM 168 188  
 FT DOMAIN 189 305  
 FT TRANSSEM 306 323  
 FT DOMAIN 324 349  
 FT TRANSSEM 350 367  
 FT DOMAIN 368 447  
 FT TRANSSEM 448 470  
 FT DOMAIN 471 539  
 FT TRANSSEM 540 560  
 FT DOMAIN 561 575  
 FT TRANSSEM 576 596  
 FT DOMAIN 597 1000  
 FT TRANSSEM 1001 1020  
 FT DOMAIN 1021 1062  
 FT TRANSSEM 1063 1076  
 FT DOMAIN 1077 1153  
 FT TRANSSEM 1154 1173  
 FT DOMAIN 1174 1275  
 FT TRANSSEM 1276 1297  
 FT DOMAIN 1298 1580  
 FT NP\_BIND 712 719  
 FT NP\_BIND 1377 1384  
 FT CARBOHYD 9  
 FT VARSPIC 739 739  
 FT VARIANT 717 717  
 SQ SEQUENCE 1580 AA; 176876 MW; B65F70A2 CRC32;

Query Match 65.3%; Score 32; DB 1; Length 1580;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KFLKFLKFL 10  
 | :||:|  
 Db 149 KTIKFKFL 157

RESULT 14  
 SY11\_BACSU  
 ID SY11\_BACSU STANDARD; PRT; 422 AA.  
 AC P22326;  
 DT 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE TYROSYL-TRNA SYNTHETASE 1 (EC 6.1.1.1) (TYROSINE--TRNA LIGASE)  
 DE (TYRS 1).  
 GN TYRS.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92138624.  
 RA HENKIN T.M., GLASS B.L., GRUNDY F.J.;  
 RT "Analysis of the Bacillus subtilis tyrs gene: conservation of a  
 regulatory sequence in multiple trna synthetase genes.";

J. BACTERIOL. 174:1299-1306(1992).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE; 98048467.  
 RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH S.D.;  
 RT "Sequencing and functional annotation of the Bacillus subtilis genes  
 in the 200 kb trnB-dnaB region.";  
 RL MICROBIOLOGY 143:3431-3441(1997).  
 [3]  
 RN SEQUENCE OF 328-422 FROM N.A.  
 RP STRAIN-168;  
 RX MEDLINE; 91035248.  
 RA GRUNDY F., HENKIN T.M.;  
 RT "Cloning and analysis of the Bacillus subtilis rpsD gene, encoding  
 ribosomal protein S4.";  
 RL J. BACTERIOL. 172:6372-6379(1990).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-TYROSINE + TRNA(TYR) -> AMP +  
 PYROPHOSPHATE + L-TYROSYL-TRNA(TYR).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING VEGETATIVE GROWTH.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M77668; G143795; --  
 DR EMBL; M59358; G143468; --  
 DR EMBL; AF008220; G2293225; --  
 DR EMBL; L17309; G861176; --  
 DR EMBL; Z99119; E1185840; --  
 DR PIR; A42648; A42648.  
 DR SUBTILIST; BG10371; TYRS.  
 DR PROSITE; PS00178; AA-TRNA\_LIGASE\_I; 1.  
 DR PFAM; PF00579; trna-synt\_1b; 1.  
 DR HSP; P00952; ITYA.  
 KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING.  
 FT SIMILAR 40 49  
 FT SIMILAR 232 236  
 FT BINDING 235 235  
 FT BINDING 235 235  
 SQ SEQUENCE 422 AA; 47737 MW; 32F5366D CRC32;

Query Match 65.3%; Score 32; DB 1; Length 422;  
 Best Local Similarity 50.0%; Pred. No. 52;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IKFLKFLKFL 10  
 | :||:|  
 Db 266 VKYLKFTFL 275

RESULT 15  
 VGNB\_CPMV  
 ID VGNB\_CPMV STANDARD; PRT; 1866 AA.  
 AC P03600;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GENOME POLYPROTEIN B [CONTAINS: PROTEASE COFACTOR; MEMBRANE BINDING  
 PROTEIN; VPG; PROTEASE (EC 3.4.22.-); RNA POLYMERASE (EC 2.7.7.48)].  
 DE COMPEA MOSAIC VIRUS (CPMV).  
 OS VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; COMOVIRIDAE;  
 OC COMOVIRUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LOMONOSOFF G.P., SHANKS M.;  
 RT "The nucleotide sequence of cowpea mosaic virus B RNA.";

RL EMBO J. 2:2253-2258(1983).  
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X00206; G58909; -  
DR PIR; A04211; GME2C.  
DR PFAM; PF00680; RNA\_dep\_RNA\_pol: 1.  
DR PFAM; PF00910; RNA\_helicase; 1.  
KW POLYPROTEIN; TRANSMEMBRANE; HYDROLASE; PROTEASE; TRANSFERASE;  
KW RNA-DIRECTED RNA POLYMERASE; ATP-BINDING.  
FT CHAIN 1 326  
FT CHAIN 327 919 PROTEASE COFACTOR (POTENTIAL).  
FT CHAIN 920 947 MEMBRANE-BINDING PROTEIN (POTENTIAL).  
FT CHAIN 948 1155 VPG PROTEIN (POTENTIAL).  
FT CHAIN 1156 1866 PROTEASE (POTENTIAL).  
FT NP\_BIND 494 501 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).  
FT ACT\_SITE 987 987 ATP (POTENTIAL).  
FT ACT\_SITE 1023 1023 THIOL PROTEASE (POTENTIAL).  
FT ACT\_SITE 1113 1113 THIOL PROTEASE (POTENTIAL).  
SQ SEQUENCE 1866 AA; 209809 MW; 75077012 CRC32;

Query Match 65.3%; Score 32; DB 1; Length 1866;  
Best Local Similarity 55.6%; Pred. No. 2.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
OY 2 KFLKFLKFL 10  
DB 1389 KFLNVRFI 1397  
||| |::|

Search completed: September 8, 1999, 00:42:59  
Job time: 2614 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:23:09 ; Search time 31.14 Seconds  
(without alignments)  
12.866 Million cell updates/sec

Title: US-09-124-280A-41

Perfect score: 49

Sequence: 1 IKFLKFLKFL 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database :

PIR\_60:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	75.5	201	2 G70130	conserved hypothet
2	36	73.5	956	2 A45506	H+-transporting AT
3	36	73.5	956	2 A43637	H+-transporting AT
4	36	73.5	956	2 S50751	H+-transporting AT
5	36	73.5	951	2 S52728	H+-transporting AT
6	36	73.5	949	2 T02083	H+-transporting AT
7	35	71.4	1278	2 A71609	probable secreted
8	34	69.4	578	2 S55102	PS2 protein - yea
9	34	69.4	564	2 E70394	ABC transporter (h
10	33	67.3	1858	1 A44214	genome polypeptide
11	33	67.3	394	2 F69888	alanine racemase h
12	33	67.3	4572	2 S57908	hypothetical polyp
13	33	67.3	110	2 S55208	hypothetical prote
14	33	67.3	1122	2 S64443	probable membrane
15	33	67.3	337	2 S33846	ribosomal protein
16	33	67.3	141	2 B70485	hypothetical prote
17	33	67.3	644	2 B70420	NADH dehydrogenase
18	32.5	66.3	238	1 E64304	probable phosphos
19	32	65.3	422	1 A42648	tyrosine--trna lig
20	32	65.3	1866	1 GWE2C	genome polypeptide
21	32	65.3	190	2 C39746	NADH dehydrogenase
22	32	65.3	201	2 S62757	NADH dehydrogenase
23	32	65.3	236	2 A56509	peroxisomal membra
24	32	65.3	260	2 A59068	2-hydroxyhepta-2,4
25	32	65.3	535	2 A64697	conserved hypothet
26	32	65.3	475	2 H71660	nitrogen assimilati
27	32	65.3	129	2 G93326	hypothetical prote
28	32	65.3	551	2 A57189	secY protein homol
29	32	65.3	553	2 T01416	secY protein homol
30	32	65.3	343	2 T01432	secY protein homol
31	32	65.3	943	2 S59317	DIP2 protein - yea
32	32	65.3	342	2 S33056	probable membrane
33	32	65.3	130	2 C70427	hypothetical prote
34	32	65.3	132	2 C71095	hypothetical prote
35	32	65.3	1113	3 JE0315	low-density lipopr
36	31	63.3	192	2 S16572	plastoquinol--plas
37	31	63.3	293	2 C70160	ribosomal protein
38	31	63.3	729	2 JQ1898	probable RNA-direc
39	31	63.3	829	2 S50246	SIP4 protein - yea

40 31 63.3 309 2 B38545 hypothetical prote  
41 31 63.3 2291 2 S11238 polymerase - Berne  
42 31 63.3 762 2 E64597 DNA mismatch repai  
43 31 63.3 762 2 H71916 DNA mismatch repai  
44 31 63.3 1279 2 E64709 type IIS restricti  
45 31 63.3 271 2 A36893 transcription acti

ALIGNMENTS

RESULT 1

G70130

conserved hypothetical protein BB0247 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Dec-1998

C:Accession: G70130

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943

A:Accession: G70130

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-201 <KLE>

A:Cross-references: GB:AF001135; GB:AE000783; NID:g2688144; PID:g2688151; TIGR:BB0247

A:Experimental source: strain B31

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 75.5%; Score 37; DB 2; Length 201;

Best Local Similarity 80.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 IKFLKFLKFL 10

| | | | |

Db 185 IAFLLKFLKFL 194

RESULT 2

A45506

H+-transporting ATPase (EC 3.6.1.35) LHA1 - tomato

C:Species: Lycopersicon esculentum (tomato)

C>Date: 14-May-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Mar-1998

C:Accession: A45506

R:Ewing, N.N.; Wimmers, L.E.; Meyer, D.J.; Chetelat, R.T.; Bennett, A.B.

Plant Physiol. 94, 1874-1881, 1990

A:Title: Molecular cloning of tomato plasma membrane H+-ATPase.

A:Reference number: A45506

A:Accession: A45506

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-956 <EWI>

A:Cross-references: GB:M60166; NID:g170463; PID:g170464

C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma

C:Keywords: ATP; hydrolase; phosphoprotein; transmembrane protein

F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>

F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted

F:427/Binding site: ATP (Lys) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 956;

Best Local Similarity 77.8%; Pred. No. 39;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KFLKFLKFL 10

| | | | |

Db 61 KFLKFLGFM 69

```
RESULT 3
A43637
H+-transporting ATPase (EC 3.6.1.35) - curled-leaved tobacco
C:Species: Nicotiana glauca (curled-leaved tobacco)
C:Date: 11-Feb-1993 #sequence_revision 12-Mar-1993 #text_change 17-Mar-1999
C:Accession: A43637
R:Botry, M.; Michelet, B.; Goffeau, A.
Biochem. Biophys. Res. Commun. 162, 567-574, 1989
A:Title: Molecular cloning of a family of plant genes encoding a protein homologous to P
A:Reference number: A43637; MUID:89334860
A:Accession: A43637
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-956 <BOU>
A:Cross-references: GB:M27888; NID:q170205; PID:q170206
A:Note: the authors translated the codon CGG for residue 131 as Trp
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:427/Binding site: ATP (Lys) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 956;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
||||| |
DB 61 KFLKFLGFM 69

RESULT 4
S50751
H+-transporting ATPase (EC 3.6.1.35) (clone PHA1) - potato
C:Species: Solanum tuberosum (potato)
C:Date: 14-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 24-Apr-1998
C:Accession: S50751; S39313
R:Harms, K.; Woehner, R.V.; Schulz, B.; Frommer, W.B.
Plant Mol. Biol. 26, 979-988, 1994
A:Title: Isolation and characterization of P-type H(+)-ATPase genes from potato.
A:Reference number: S50751
A:Accession: S50751
A:Molecule type: mRNA
A:Residues: 1-956 <HAR>
A:Cross-references: EMBL:X76536; NID:g435002; PID:g435003
A:Note: the nucleotide is not given
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:485-657/Domain: ATPase nucleotide-binding domain homology <ATN>
F:333/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:427/Binding site: ATP (Lys) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 956;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
||||| |
DB 61 KFLKFLGFM 69

RESULT 5
S52728
H+-transporting ATPase (EC 3.6.1.35) - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 24-Apr-1998
C:Accession: S52728
R:Campos, F.; Perez-Castineira, J.R.; Villalba, J.M.; Serrano, R.
submitted to the EMBL Data Library, March 1995
A:Reference number: S52728
A:Accession: S52728
```

```
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-951 <CAM>
A:Cross-references: EMBL:X85804; NID:g758249; PID:g758250
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: ATP; hydrolase; phosphoprotein
F:483-655/Domain: ATPase nucleotide-binding domain homology <ATN>
F:331/Active site: Asp (aspartylphosphate intermediate) #status predicted
F:425/Binding site: ATP (Lys) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 951;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
||||| |
DB 59 KFLKFLGFM 67

RESULT 6
T02083
H+-transporting ATPase (EC 3.6.1.35) Mhal - maize
C:Species: Zea mays (maize)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 26-Feb-1999
C:Accession: T02083
R:Jin, Y.; Bennetzen, J.L.
Plant Cell 6, 1177-1186, 1994
A:Title: Integration and nonrandom mutation of a plasma membrane proton ATPase gene f
A:Reference number: Z14555
A:Accession: T02083
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-949 <JIN>
A:Cross-references: EMBL:U09989; NID:g507770; PID:g533775
A:Experimental source: strain D3L
C:Genetics:
A:Gene: Mhal
A:Map position: 2L
A:Introns: 21/3; 51/3; 94/3; 140/3; 185/3; 245/3; 314/3; 354/3; 435/3; 470/3;
C:Function:
A:Description: transports protons across the plasma membrane to regulate intracellular
C:Keywords: ATP; hydrogen ion transport; hydrolase; phosphoprotein

Query Match 73.5%; Score 36; DB 2; Length 949;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
||||| |
DB 64 KFLKFLGFM 72

RESULT 7
A71609
probable secreted protein PFB0675w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: A71609
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: A71609
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1278 <GAR>
A:Cross-references: GB:AF001410; GB:AE001362; NID:g3845245; PID:g3845247; TIGR:PFB067
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0675w
```



Query Match 71.4%; Score 35; DB 2; Length 1278;  
Best Local Similarity 70.0%; Pred. No. 78;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10  
||| ||| |  
Db 687 IKFCKFLFLFI 696

RESULT 8  
S55102  
PSP2 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YM9571.01; protein YML017w  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 06-Feb-1998  
C:Accession: S55102; S59358  
R:Gentiles, S.; Bowman, S.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S55102  
A:Accession: S55102  
A:Molecule type: DNA  
A:Residues: 1-578 <GEN>  
A:Cross-references: EMBL:Z49810; NID:9854472; PID:9854473; MIPS:YML017w  
A:Experimental source: strain AB972  
R:Formosa, T.D.; Nittis, T.  
submitted to the EMBL Data Library, August 1995  
A:Description: High copy suppressors of the temperature sensitivity of DNA polymerase alpha  
A:Reference number: S59358  
A:Accession: S59358  
A:Molecule type: DNA  
A:Residues: 1-578 <FOR>  
A:Cross-references: EMBL:U33116; NID:9995752; PID:9995753  
C:Genetics:  
A:Gene: SGD:PS2  
A:Cross-references: SGD:S0004479; MIPS:YML017w  
A:Map position: 13L

Query Match 69.4%; Score 34; DB 2; Length 578;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKF 9  
|| ||| |  
Db 117 KETKFKF 124

RESULT 9  
E70394  
ABC transporter (hlyB subfamily) - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 18-Sep-1998  
C:Accession: E70394  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt, V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: E70394  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-564 <AQ>  
A:Cross-references: GB:AE000723; NID:92983569; PID:92983577; GB:AE000657  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: abcT5  
C:Superfamily: ATP-binding cassette homology  
F:349-543/Domain: ATP-binding cassette homology <ABC>  
F:366-373/Region: nucleotide-binding motif A (P-loop)

Query Match 69.4%; Score 34; DB 2; Length 564;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKF 9  
:|||||:  
Db 310 EFLKFLRF 317

RESULT 10  
A44214  
genome polyprotein 1 - cowpea severe mosaic virus  
N:Contains: 24k protein; 32k protein; 58k membrane-binding protein; RNA-directed RNA  
C:Species: cowpea severe mosaic virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 31-Dec-1996  
C:Accession: A44214  
R:Chen, X.; Bruening, G.  
Virology 191, 607-618, 1992  
A:Title: Cloned DNA copies of cowpea severe mosaic virus genomic RNAs: infectious tra  
A:Reference number: A44214; MUID:93079863  
A:Accession: A44214  
A:Molecule type: genomic RNA  
A:Residues: 1-1858 <CHE>  
A:Cross-references: GB:M83830  
C:Genetics:  
A:Map position: segment 1  
C:Superfamily: cowpea aphid-borne mosaic virus genome polyprotein B  
C:Keywords: genome-linked protein; glycoprotein; hydrolase; nucleotidyltransferase; p  
F:2-313/Product: 32k proteinase cofactor #status predicted <PCF>  
F:218-235/Region: hydrophobic  
F:314-908/Product: 58k membrane-binding protein #status predicted <MBP>  
F:891-907/Domain: transmembrane #status predicted <TMN>  
F:909-936/Product: genome-linked protein Vp #status predicted <VPG>  
F:937-1146/Product: 24k viral proteinase #status predicted <VPT>  
F:1147-1858/Product: RNA-directed RNA polymerase #status predicted <RRP>  
F:1811-1826/Region: hydrophobic  
F:326,958,1744,1758/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.3%; Score 33; DB 1; Length 1858;  
Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10  
|||||:  
Db 1379 KFLKFLVFI 1387

RESULT 11  
F69888  
alanine racemase homolog yncD - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1998  
C:Accession: F69888  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segkora, P.; Tognoni, A.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil  
A:Reference number: A69580; MUID:98044033  
A:Accession: F69888  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-394 <KUN>  
 A:Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PID:ell83423; PID:g2634148  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yncd  
 C:Superfamily: alanine racemase

Query Match 67.3%; Score 33; DB 2; Length 394;  
 Best Local Similarity 66.7%; Pred. No. 57;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLKF 10  
 ||: ||| ||  
 Db 187 KFISFLSFL 195

RESULT 12  
 S57908  
 hypothetical polyprotein - rice  
 C:Species: Oryza sativa (rice)  
 C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: S57908; S33529  
 R:Moriyama, H.; Nitta, T.; Fukuhara, T.  
 Mol. Gen. Genet. 248, 364-369, 1995  
 A:Title: Double-stranded RNA in rice: a novel RNA replicon in plants.  
 A:Reference number: S57908; MUID:96004703  
 A:Accession: S57908  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-4572 <MOR>  
 A:Cross-references: EMBL:D32136; NID:g511666; PID:dl007437; PID:gl030716  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 A:Note: only a part of the coding sequence is given  
 R:Fukuhara, T.; Moriyama, H.; Pak, J.I.; Hyakutake, H.; Nitta, T.  
 Plant Mol. Biol. 21, 1121-1130, 1993  
 A:Title: Enigmatic double-stranded RNA in Japonica rice.  
 A:Reference number: S33529; MUID:93257627  
 A:Accession: S33529  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 3854-4370 <FUK>  
 A:Cross-references: EMBL:D12639  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992  
 C:Keywords: polypeptide

Query Match 67.3%; Score 33; DB 2; Length 4572;  
 Best Local Similarity 87.5%; Pred. No. 6e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KFLKFLKF 9  
 ||||| |||  
 Db 188 KFLKTLKF 195

RESULT 13  
 S55208  
 hypothetical protein YJR020w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein J1458; hypothetical protein YJR03.15  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 12-Dec-1997  
 C:Accession: S55208; S57035  
 R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: S55183  
 A:Accession: S55208  
 A:Molecule type: DNA  
 A:Residues: 1-110 <DEH>  
 A:Cross-references: EMBL:X87611; NID:g854567; PID:g854593  
 R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.  
 submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S56771

A:Accession: S57035  
 A:Molecule type: DNA  
 A:Residues: 1-110 <ZAG>  
 A:Cross-references: EMBL:Z49519; NID:g1015654; PID:g1015656; MIPS:YJR020w  
 C:Genetics:  
 A:Map position: 10R

Query Match 67.3%; Score 33; DB 2; Length 110;  
 Best Local Similarity 66.7%; Pred. No. 17;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKFLKFLKF 9  
 :||| ||:|  
 Db 1 MKELPFLRF 9

RESULT 14  
 S64443  
 probable membrane protein YGR134w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein G6398  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 14-Nov-1997  
 C:Accession: S64443  
 R:Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Tallia, E.; Nawrocki, A.;  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64448  
 A:Accession: S64443  
 A:Molecule type: DNA  
 A:Residues: 1-1122 <VAN>  
 A:Cross-references: EMBL:Z72919; NID:g1323222; PID:e243526; PID:g1323223; MIPS:YGR134  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Map position: 7R  
 C:Keywords: transmembrane protein  
 F:65-81/Domain: transmembrane #status predicted <TM1>  
 F:154-170/Domain: transmembrane #status predicted <TM2>  
 F:463-479/Domain: transmembrane #status predicted <TM3>  
 F:955-971/Domain: transmembrane #status predicted <TM4>

Query Match 67.3%; Score 33; DB 2; Length 1122;  
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IKFLKFLKF 10  
 | ||||| ||  
 Db 125 ILLKFLRFL 134

RESULT 15  
 S53846  
 ribosomal protein S7 - Acanthamoeba castellanii mitochondrion (SGC6)  
 C:Species: mitochondrion Acanthamoeba castellanii  
 C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Feb-1999  
 C:Accession: S53846  
 R:Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.  
 J. Mol. Biol. 245, 522-537, 1995  
 A:Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: c  
 A:Reference number: S53825; MUID:95147275  
 A:Accession: S53846  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-337 <BUR>  
 A:Cross-references: GB:U12386; NID:g562028; PID:g562050  
 A:Experimental source: strain Neff; ATCC 30010  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC6  
 C:Keywords: mitochondrion

Query Match 57.38; Score 33; DB 2; Length 337;  
Best Local Similarity 75.0%; Pred. No. 49;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IKFLKFLK 8  
|:|:|:|  
Db 327 IRFLRFLK 334

Search completed: September 7, 1999, 23:54:50  
Job time: 1901 sec

**THIS PAGE BLANK (USPTO)**

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	49	100.0	11	1	US-08-049-871-8	Sequence 8, Appl
2	49	100.0	11	1	US-07-819-893-8	Sequence 8, Appl
3	49	100.0	11	1	US-08-280-397-8	Sequence 8, Appl
4	49	100.0	11	1	US-08-456-1128-37	Sequence 37, Appl
5	49	100.0	10	2	US-08-456-1128-41	Sequence 41, Appl
6	45	91.8	10	1	US-08-097-830E-8	Sequence 8, Appl
7	45	91.8	12	1	US-08-097-830E-20	Sequence 20, Appl
8	45	91.8	10	2	US-08-456-1128-8	Sequence 8, Appl
9	45	91.8	12	2	US-08-456-1128-20	Sequence 20, Appl
10	35	71.4	7	1	US-07-819-893-9	Sequence 9, Appl
11	35	71.4	7	1	US-08-280-397-9	Sequence 9, Appl
12	35	71.4	7	1	US-08-097-830E-7	Sequence 7, Appl
13	35	71.4	7	2	US-08-456-1128-7	Sequence 7, Appl
14	35	71.4	7	2	US-08-456-1128-38	Sequence 38, Appl
15	33	67.3	10	1	US-08-097-830E-10	Sequence 10, Appl
16	33	67.3	10	2	US-08-456-1128-10	Sequence 10, Appl
17	33	67.3	9	2	US-08-456-1128-40	Sequence 40, Appl
18	31	63.3	254	2	US-08-286-819A-28	Sequence 28, Appl
19	30	61.2	287	1	US-08-457-245-3	Sequence 3, Appl
20	30	61.2	185	2	US-08-338-543-2	Sequence 2, Appl
21	30	61.2	576	2	US-08-676-279-58	Sequence 58, Appl
22	29	59.2	135	1	US-08-468-853-4	Sequence 4, Appl
23	29	59.2	141	1	US-08-363-300-2	Sequence 2, Appl
24	29	59.2	135	2	US-08-468-855-4	Sequence 4, Appl
25	29	59.2	627	2	US-08-703-947-2	Sequence 2, Appl
26	29	59.2	135	2	US-08-310-357-4	Sequence 4, Appl
27	29	59.2	135	2	US-08-468-852-4	Sequence 4, Appl
28	29	59.2	238	2	US-08-380-403A-7	Sequence 7, Appl
29	28	57.1	887	1	US-08-215-709-1	Sequence 1, Appl
30	28	57.1	32	1	US-08-118-270-188	Sequence 188, App
31	28	57.1	308	1	US-08-457-245-8	Sequence 8, Appl
32	28	57.1	350	1	US-08-415-751-16	Sequence 16, Appl
33	28	57.1	440	1	US-08-307-499-15	Sequence 15, Appl
34	28	57.1	116	1	US-08-702-344-21	Sequence 21, Appl
35	28	57.1	51	1	US-08-691-641-3	Sequence 3, Appl
36	28	57.1	3056	2	US-08-508-836A-8	Sequence 8, Appl
37	28	57.1	3056	2	US-08-629-001A-3	Sequence 3, Appl
38	28	57.1	1581	2	US-08-404-531B-6	Sequence 6, Appl
39	28	57.1	1582	2	US-08-404-531B-9	Sequence 9, Appl

;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson, Costigan & Hoare  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/819,893  
;; FILING DATE: 19920115  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: circular  
;; US-07-819-893-8

Query Match 100.0%; Score 49; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
Db 1 IKFLKFLKFL 10

RESULT 3  
US-08-280-397-8  
;; Sequence 8, Application US/08280397  
;; Patent No. 5589459  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
;; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
;; TITLE OF INVENTION: Prevention and Treatment of Septic  
;; TITLE OF INVENTION: Shock  
;; NUMBER OF SEQUENCES: 10  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: IBM PS/2  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/280,397  
;; FILING DATE: 07/28/94  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/819,893  
;; FILING DATE: 01/16/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-280-397-8

Query Match 100.0%; Score 49; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
Db 1 IKFLKFLKFL 10

RESULT 4  
US-08-456-112B-37  
;; Sequence 37, Application US/08456112B  
;; Patent No. 5834430  
;; GENERAL INFORMATION:  
;; APPLICANT: Porro, Massimo  
;; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hedman, Gibson & Costigan  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
;; COMPUTER: LEADING EDGE 486  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,112B  
;; FILING DATE: May 31, 1995  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-004  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: circular  
;; US-08-456-112B-37

Query Match 100.0%; Score 49; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10  
1111111111  
Db 1 IKFLKFLKFL 10

## RESULT 5

US-08-456-112B-41  
; Sequence 41, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 41:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-41

Query Match 100.0%; Score 49; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0024;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10  
1111111111  
Db 1 IKFLKFLKFL 10

## RESULT 6

US-08-097-830E-8  
; Sequence 8, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-8

Query Match 91.8%; Score 45; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10  
1111111111  
Db 1 KFLKFLKFL 9

## RESULT 7

US-08-097-830E-20  
; Sequence 20, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids

```

; TYPE: amino acid
; TOPOLOGY: circular
; FEATURE:
; OTHER INFORMATION: sulfide bond between Cys and Cys
US-08-097-830E-20

```

```

Query Match          91.8%; Score 45; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 KFLKFLKFL 10
   |||||
Db 2 KFLKFLKFL 10

```

```

RESULT 8
US-08-456-112B-8
; Sequence 8, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:

```

```

; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-8

```

```

Query Match          91.8%; Score 45; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 KFLKFLKFL 10
   |||||
Db 1 KFLKFLKFL 9

```

```

RESULT 9
US-08-456-112B-20
; Sequence 20, Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo

```

```

; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
US-08-456-112B-20

```

```

Query Match          91.8%; Score 45; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 KFLKFLKFL 10
   |||||
Db 2 KFLKFLKFL 10

```

```

RESULT 10
US-07-819-893-9
; Sequence 9, Application US/07819893
; Patent No. 5371186
; GENERAL INFORMATION:
; APPLICANT: Porro, Massimo
; TITLE OF INVENTION: Synthetic Peptides for Detoxification
; TITLE OF INVENTION: of Bacterial Endotoxins and for the
; TITLE OF INVENTION: Prevention and Treatment of Septic
; TITLE OF INVENTION: Shock
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson, Costigan & Hoare
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2 DOS
; OPERATING SYSTEM:
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/819,893
; FILING DATE: 19920115
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

```



;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Costigan, James V.  
;; REGISTRATION NUMBER: 25,669  
;; REFERENCE/DOCKET NUMBER: 576-002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 302-8989  
;; TELEFAX: (212) 302-8998  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
US-07-819-893-9

Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 11  
US-08-280-397-9  
; Sequence 9, Application US/08280397  
; Patent No. 5589459  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Synthetic Peptides for Detoxification  
; TITLE OF INVENTION: of Bacterial Endotoxins and for the  
; TITLE OF INVENTION: Prevention and Treatment of Septic  
; TITLE OF INVENTION: Shock  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/280.397  
; FILING DATE: 07/26/94  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/819,893  
; FILING DATE: 01/16/92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-280-397-9

Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 12  
US-08-097-830E-7  
; Sequence 7, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-7

Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 13  
US-08-456-112B-7  
; Sequence 7, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA

; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-7

Query Match 71.4%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 14  
US-08-456-112B-38  
; Sequence 38, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-38

Query Match 71.4%; Score 35; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.9e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLK 8  
Db 1 KFLKFLK 7

RESULT 15  
US-08-097-830E-10  
; Sequence 10, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,830E  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-097-830E-10

Query Match 67.3%; Score 33; DB 1; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.3;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFLKFLK 9  
Db 1 KFFKFLK 8

Search completed: September 7, 1999, 23:23:07  
Job time: 911 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 22:50:24 ; Search time 40.95 Seconds  
(without alignments)  
5.784 Million cell updates/sec

Title: US-09-124-280A-41

Perfect score: 49

Sequence: 1 IKFLKFLKFL 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	11	1 R33532	Peptide for treati
2	49	100.0	11	1 R39294	Endotoxin lipid A
3	49	100.0	11	1 W21625	Antibiotic potenti
4	49	100.0	10	1 W21629	Antibiotic potenti
5	47	95.9	32	1 R33514	Generic peptide fo
6	45	91.8	10	1 R71779	Peptide neutralisi
7	45	91.8	12	1 R71791	Peptide neutralisi
8	45	91.8	10	1 W21596	Antibiotic potenti
9	45	91.8	12	1 W21608	Antibiotic potenti
10	36	73.5	420	1 W19781	Tyrosyl-tRNA synth
11	35	71.4	7	1 R33533	Peptide for treati
12	35	71.4	7	1 R39295	Endotoxin lipid A
13	35	71.4	7	1 R71778	Peptide neutralisi
14	35	71.4	7	1 W21626	Antibiotic potenti
15	35	71.4	7	1 W21595	Antibiotic potenti
16	33	67.3	10	1 R71781	Peptide neutralisi
17	33	67.3	9	1 W21628	Antibiotic potenti
18	33	67.3	10	1 W21598	Antibiotic potenti
19	32	65.3	1580	1 W57412	Homo sapiens sulph
20	31	63.3	2408	1 R24306	Translation of ORF
21	31	63.3	540	1 W20605	H. pylori cytoplas
22	31	63.3	566	1 W20945	H. pylori inner me
23	31	63.3	1279	1 W98761	H. pylori GHPO 109
24	30	61.2	6	1 R33515	Peptide for treati
25	30	61.2	185	1 R47115	Toxoplasma GP28.5
26	30	61.2	404	1 R48674	Putative peach pol
27	30	61.2	1077	1 R91597	C3G protein. Ras p
28	30	61.2	287	1 W05967	Cyclopropane mycol
29	30	61.2	470	1 W12678	DAX-1 protein. New
30	30	61.2	252	1 W18315	Toxoplasma gondii
31	30	61.2	252	1 W36012	Toxoplasma gondii
32	30	61.2	252	1 W32084	Non-glycosylated T
33	30	61.2	184	1 W32085	Non-glycosylated T
34	30	61.2	252	1 W32083	Amino acid sequenc
35	30	61.2	555	1 W84154	Human desaturase e
36	30	61.2	608	1 W84155	Human desaturase e
37	30	61.2	746	1 W84156	Human desaturase e
38	30	61.2	746	1 W85135	Human desaturase e
39	30	61.2	555	1 W85133	A desaturase enzym
40	30	61.2	608	1 W85134	A desaturase enzym
41	30	61.2	562	1 W95512	Amino acid sequenc
42	30	61.2	615	1 W95513	Amino acid sequenc
43	30	61.2	753	1 W95514	Amino acid sequenc

ALIGNMENTS

RESULT 1

R33532  
ID R33532 standard; peptide; 11 AA.  
AC R33532;  
DT 07-JUL-1993 (first entry)  
DE Peptide for treating septic shock.  
KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
KW cytokine release control; treatment; pertussis; bacterial meningitis;  
KW HIV related infections; polymyxin B.  
OS Synthetic.  
PN ZA9200943-A.  
PD 25-NOV-1992.  
PF 10-FEB-1992; 000943.  
PR 11-FEB-1991; US-658744.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-094304/11.  
PT New peptide for treatment or prevention of toxic shock - comprises  
PT specified sequences of aminoacid(s) and analogs  
PT comprising sequences retro-orientated  
PS Claim 10: page 32; 39pp: English.  
CC This peptide is a specific example of a generic peptide of  
CC formula R-( Lys/Arg/His - Phe/Tyr/Trp - Leu/Ile/Val )n-R, where  
CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
CC residue. The peptide is useful for treating or preventing septic  
CC shock, mixing with polymyxin B to reduce its toxicity; removing  
CC endotoxins from blood, sera or other fluids (in vivo or in  
CC vitro); controlling release of cytokines induced by endotoxins;  
CC as diagnostic reagents to detect and quantify toxins in blood  
CC or sera; preparing non-toxic antigenic complexes of lipid A or  
CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
CC meningitis and HIV-related infections. The usual dose is 10-100  
CC ug/kg/day, given parenterally. It binds to the same sites as  
CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
CC has no antibiotic activity; does not lyse erythrocytes; has no  
CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
CC against proteases.  
SQ Sequence 11 AA;

Query Match 100.0%; Score 49; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IKFLKFLKFL 10  
|||||  
DB 1 IKFLKFLKFL 10

RESULT 2

R39294  
ID R39294 standard; peptide; 11 AA.  
AC R39294;  
DT 22-DEC-1993 (first entry)  
DE Endotoxin lipid A neutralising peptide.  
KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
KW Polymyxin B; PMB; toxicity reduction; reduce; blood endotoxin;  
KW removal; sera; vaccines; drug soins.; solutions; cytokine release;  
KW control; in vivo; in vitro; detoxification; detection;  
OS Quantification.  
OS Synthetic.  
PN W09314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.

PI Porro M;  
 DR WPI: 93-243143/30.  
 PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
 PT forming non-toxic, antigenic complex, used to treat or prevent  
 PT septic shock, in vaccines to detoxify blood, etc.  
 PS Claim 10; Page 33; 45pp; English.  
 CC The sequence is that of a peptide which binds to the lipid A  
 CC component of bacterial endotoxin at the same site as polymyxin B  
 CC (PMB) and with about the same affinity to produce a non-toxic  
 CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
 CC proteolytic degradation in serum, has no antibiotic activity and no  
 CC haemolytic action. It is especially used to treat or prevent septic  
 CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
 CC sera, vaccines, drug solutions, etc.; to control release of cytokines  
 CC induced by endotoxins; for in vivo or in vitro detoxification of  
 CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
 CC products. 11 AA;  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 49; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 | | | | | | | | | |  
 Db 1 IKFLKFLKFL 10

RESULT 3  
 ID W21625 standard; peptide; 11 AA.  
 AC W21625;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence; reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 49; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 | | | | | | | | | |  
 Db 1 IKFLKFLKFL 10

RESULT 4  
 ID W21629 standard; peptide; 10 AA.  
 AC W21629;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #37.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence; reduces dose of antibiotic required  
 PS Claim 41; Page 28; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 49; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 | | | | | | | | | |  
 Db 1 IKFLKFLKFL 10

RESULT 5  
 ID R33514 standard; peptide; 32 AA.  
 AC R33514;  
 DT 07-JUL-1993 (first entry)  
 DE Generic peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #41.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 45; Page 29; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 49; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10  
 | | | | | | | | | |  
 Db 1 IKFLKFLKFL 10

RESULT 5  
 ID R33514 standard; peptide; 32 AA.  
 AC R33514;  
 DT 07-JUL-1993 (first entry)  
 DE Generic peptide for treating septic shock.  
 KW Toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI: 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 2; Page 32; 39pp; English.  
 CC This is the longest sequence of a generic peptide X-(Lys-Phe-Leu)n-X  
 CC where n = 1-10, the peptide is useful for treating or preventing  
 CC septic shock, mixing with polymyxin B to reduce its toxicity;  
 CC removing endotoxins from blood, sera or other fluids (in vivo  
 CC or in vitro); controlling release of cytokines induced by  
 CC endotoxins; as diagnostic reagents to detect and quantify toxins  
 CC in blood or sera; preparing non-toxic antigenic complexes of  
 CC lipid A or lipopolysaccharide (LPS); and for treating pertussis,  
 CC bacterial meningitis and HIV-related infections. The usual dose is  
 CC 10-100 ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 SQ Sequence 32 AA;

```

Query Match          95.9%; Score 47; DB 1; Length 32;
Best Local Similarity 90.0%; Pred. No. 0.014;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKFLKFLKFL 10
DB 4 LKFLKFLKFL 13

RESULT 6
R71779
ID R71779 standard; peptide; 10 AA.
AC R71779;
DT 01-OCT-1995 (first entry)
DE Peptide neutralising toxicity of Lipid A.
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.
OS Synthetic.
PN WQ9503327-A.
PD 02-FEB-1995.
PF 21-JUL-1994; E02413.
PR 26-JUL-1993; US-097830.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M.
DR WPI; 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising
PT repeating units of a basic aminoacid or basic and hydrophobic
PT amino acids
PS Claim 10; Page 20; 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides of formula:
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or
CC greater.
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic
CC activity. Hence they can be used therapeutically to treat septic shock
CC and also in vitro to detoxify vaccines, drug solutions, injectable
CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new
CC peptides. 10 AA;
SQ Sequence 10 AA;

Query Match          91.8%; Score 45; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
DB 1 KFLKFLKFL 9

RESULT 7
R71791
ID R71791 standard; peptide; 12 AA.
AC R71791;
DT 01-OCT-1995 (first entry)
DE Peptide neutralising toxicity of Lipid A.
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.
OS Synthetic.
FH Key Location/Qualifiers
FT disulfide bond 1..12
PN WQ9503327-A.
PD 02-FEB-1995.
PF 21-JUL-1994; E02413.
PR 26-JUL-1993; US-097830.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M.
DR WPI; 95-075190/10.
PT New peptide(s) for neutralising LPS endotoxin - comprising
PT repeating units of a basic aminoacid or basic and hydrophobic
PT amino acids

```

```

PS Claim 22; Page 21; 26pp; English.
CC New peptides are claimed which are linear or cyclic peptides which
CC include units of formula: (A)n, where A is the cationic amino acid Lys
CC or Arg and n is 7-10; (AB)m, where A is Lys or Arg, B is the hydrophobic
CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and
CC (ABC)p, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or
CC Trp, and p is 2 or greater.
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic
CC activity. Hence they can be used therapeutically to treat septic shock
CC and also in vitro to detoxify vaccines, drug solutions, injectable
CC nutrient solutions, etc.
CC The present sequence is a specifically claimed example of the new
CC peptides. 12 AA;
SQ Sequence 12 AA;

Query Match          91.8%; Score 45; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
DB 2 KFLKFLKFL 10

RESULT 8
W21596
ID W21596 standard; peptide; 10 AA.
AC W21596;
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #8.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.
OS Synthetic.
PN WQ9638163-A1.
PD 05-DEC-1996.
PF 29-MAY-1996; E02313.
PR 31-MAY-1995; US-456112.
PA (BIOS-) BIOSYNTH SRL.
PI Porro M, Varra M.
DR WPI; 97-034095/03.
PT Potentiating activity of antibiotic with peptide contg. cationic
PT amino acid sequence - reduces dose of antibiotic required
PS Claim 12; Page 24; 37pp; English.
CC The sequences given in W21589-633 represent peptides which act to
CC potentiate the activity of an antibiotic when they are co-administered
CC with the antibiotic. Compositions containing these peptides are used
CC to treat or prevent microbial infections. These peptides bind to
CC lipopolysaccharide on the bacteria so may increase permeability of
CC the outer bacterial membrane to the antibiotic, allowing a reduction
CC in the dose of antibiotic required by 10-90% of the normal dose for
CC in vivo or in vitro application. Any toxic side effects are
CC correspondingly reduced.
SQ Sequence 10 AA;

Query Match          91.8%; Score 45; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFLKFLKFL 10
DB 1 KFLKFLKFL 9

RESULT 9
W21608
ID W21608 standard; peptide; 12 AA.
AC W21608;
DT 26-AUG-1997 (first entry)
DE Antibiotic potentiating peptide #20.
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;
KW permeability; outer bacterial membrane.

```

OS Synthetic.  
 FH Key Location/Qualifiers  
 PN disulfide bond 1. 12  
 PD WO9638163-A1.  
 PF 05-DEC-1996.  
 PR 29-MAY-1996; E02313.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI; 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PS amino acid sequence - reduces dose of antibiotic required  
 CC Claim 24; Page 26; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 CC Sequence 12 AA;  
 SQ

Query Match 91.8%; Score 45; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLKFL 10  
 Db 2 KFLKFLKFL 10

RESULT 10  
 W19781 ID W19781 standard; Protein; 420 AA.  
 AC 10-SEP-1997 (first entry)  
 DE Tyrosyl-tRNA synthetase from Staph. aureus.  
 KW antibody; Bacillus subtilis; tyrs; immunological response;  
 KW vaccine; skin; protection.  
 OS Staphylococcus aureus.  
 PN EP-785258-A1.  
 PD 23-JUL-1997.  
 PF 17-JAN-1997; 300307.  
 PR 19-JAN-1996; GB-001067.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PI Hodgson JE, Lawlor EJ;  
 DR WPI; 97-365933/34.  
 DR N-PSDB; 072807.  
 PT DNA encoding tyrosyl-tRNA synthetase from Staphylococcus aureus WCUH  
 PT 29 - useful for protection against bacterial infections  
 PS Claim 14; Page 19-21; 31pp; English.  
 CC The present sequence is a novel tyrosyl tRNA synthetase protein from  
 CC Staphylococcus aureus strain WCUH29 (NCIMB 40771), which is related by  
 CC amino acid sequence homology to Bacillus subtilis tRNA synthetase encoded  
 CC by the tyrs gene. The enzyme catalyses the aminoacylation of tRNA-Tyr, by  
 CC a two step mechanism. The first step involves formation of a stable  
 CC enzyme:tyrosinyl adenylate complex resulting from the specific binding  
 CC and reaction of ATP and L-tyrosine. Subsequently, the 3' terminal  
 CC adenosine of enzyme-bound tRNA-Tyr reacts with the aminoacyl adenylate,  
 CC leading to the esterification of the tRNA and release of AMP. Vectors,  
 CC comprising the DNA (or polynucleotides having at least 70 % identity to  
 CC it) can be used for the recombinant production of the enzyme. The enzyme  
 CC or its related DNA (through gene therapy) is used to induce an  
 CC immunological response in a mammal to generate antibodies to protect  
 CC against disease. The antibodies protect against invasion of bacteria,  
 CC e.g. by blocking adherence of bacteria to damaged tissue, including  
 CC wounds in skin or connective tissue caused by mechanical, chemical or  
 CC thermal damage or by implantation of in-dwelling devices, or wounds in  
 CC the mucous membranes.

SQ Sequence 420 AA;  
 Query Match 73.5%; Score 36; DB 1; Length 420;  
 Best Local Similarity 70.0%; Pred. No. 14;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IKFLKFLKFL 10  
 Db 265 IKFLKFLKFL 274

RESULT 11  
 R33533 ID R33533 standard; peptide; 7 AA.  
 AC R33533;  
 DT 07-JUL-1993 (first entry)  
 DE Peptide for treating septic shock.  
 KW toxic shock; blood endotoxin removal; serum; diagnostic reagent;  
 KW cytokine release control; treatment; pertussis; bacterial meningitis;  
 KW HIV related infections; polymyxin B.  
 OS Synthetic.  
 PN ZA9200943-A.  
 PD 25-NOV-1992.  
 PF 10-FEB-1992; 000943.  
 PR 11-FEB-1991; US-658744.  
 PA (PORR/) PORRO M.  
 PI Porro M;  
 DR WPI; 93-094304/11.  
 PT New peptide for treatment or prevention of toxic shock - comprises  
 PT specified sequences of aminoacid(s) and analogs  
 PT comprising sequences retro-orientated  
 PS Claim 11; Page 32; 39pp; English.  
 CC This peptide is a specific example of a generic peptide of  
 CC formula R-( Lys/Arg/His - Phe/Tyr/Irp - Leu/Ile/Val )n-R, where  
 CC n = 1-100 and each R is H, an amino acid residue or a fatty acid  
 CC residue. The peptide is useful for treating or preventing septic  
 CC shock, mixing with polymyxin B to reduce its toxicity; removing  
 CC endotoxins from blood, sera or other fluids (in vivo or in  
 CC vitro); controlling release of cytokines induced by endotoxins;  
 CC as diagnostic reagents to detect and quantify toxins in blood  
 CC or sera; preparing non-toxic antigenic complexes of lipid A or  
 CC lipopolysaccharide (LPS); and for treating pertussis, bacterial  
 CC meningitis and HIV-related infections. The usual dose is 10-100  
 CC ug/kg/day, given parenterally. It binds to the same sites as  
 CC polymyxin B, i.e. it inhibits all the toxic effects of lipid A. It  
 CC has no antibiotic activity; does not lyse erythrocytes; has no  
 CC toxicity in mice when injected at 50mg/kg and is relatively unstable  
 CC against proteases.  
 CC Sequence 7 AA;  
 SQ

Query Match 71.4%; Score 35; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLK 8  
 Db 1 KFLKFLK 7

RESULT 12  
 R39295 ID R39295 standard; peptide; 7 AA.  
 AC R39295;  
 DT 22-DEC-1993 (first entry)  
 DE Endotoxin lipid A neutralising peptide.  
 KW Bacterial endotoxin; non-toxic; treatment; prevention; toxic shock;  
 KW polymyxin B; PWB; toxicity reduction; reduce; blood endotoxin;  
 KW removal; sera; vaccines; drug solns.; solutions; cytokine release;  
 KW control; in vivo; in vitro; detoxification; detection;  
 OS Synthetic.

PN W09314115-A.  
PD 22-JUL-1993.  
PF 14-MAY-1992; E01060.  
PR 16-JAN-1992; US-819893.  
PA (PORR/) PORRO M.  
PI Porro M;  
DR WPI; 93-243143/30.  
PT New peptide(s) which neutralise lipid A of bacterial endotoxin -  
PT forming non-toxic, antigenic complex, used to treat or prevent  
PT septic shock, in vaccines to detoxify blood, etc.  
PS Claim 11; Page 33; 45pp; English.  
CC The sequence is that of a peptide which binds to the lipid A  
CC component of bacterial endotoxin at the same site as polymyxin B  
CC (PMB) and with about the same affinity to produce a non-toxic  
CC antigenic complex. Unlike PMB it is not toxic, is susceptible to  
CC proteolytic degradation in serum, has no antibiotic activity and no  
CC haemolytic action. It is especially used to treat or prevent septic  
CC shock; to reduce toxicity of PMB; to remove endotoxins from blood,  
CC sera, vaccines, drug solns. etc.; to control release of cytokines  
CC induced by endotoxins; for in vivo or in vitro detoxification of  
CC bacterial endotoxins, and to detect or quantify endotoxins in blood  
CC products. 7 AA;  
SQ Sequence 7 AA;  
Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KFLKFLK 8  
DB 1 KFLKFLK 7  
RESULT 13  
R71778  
ID R71778 standard; peptide; 7 AA.  
AC R71778;  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
PD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic amino acid or basic and hydrophobic  
PT amino acids  
PS Claim 9; Page 20; 26pp; English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or  
CC greater.  
CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
CC activity. Hence they can be used therapeutically to treat septic shock  
CC and also in vitro to detoxify vaccines, drug solutions, injectable  
CC nutrient solutions, etc.  
CC The present sequence is a specifically claimed example of the new  
CC peptides. 7 AA;  
SQ Sequence 7 AA;  
Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KFLKFLK 8

DB 1 KFLKFLK 7  
RESULT 14  
W21626  
ID W21626 standard; peptide; 7 AA.  
AC W21626;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #38.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 42; Page 28; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 7 AA;  
Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KFLKFLK 8  
DB 1 KFLKFLK 7  
RESULT 15  
W21595  
ID W21595 standard; peptide; 7 AA.  
AC W21595;  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #7.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 11; Page 24; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 7 AA;

Wed Sep 8 09:45:04 1999

Query Match 71.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KFLKFLK 8  
      |||||  
Db 1 KFLKFLK 7

Search completed: September 7, 1999, 23:18:19  
Job time: 1675 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:17:30 ; Search time 41.46 Seconds  
(without alignments)  
14.844 Million cell updates/sec

Title: US-09-124-280A-42  
Perfect score: 50  
Sequence: 1 KKKKKKFLFL 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

- Database : SPTREMBL10.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	84.0	730	Q99248	Q99248 saccharomyc
2	41	82.0	1817	Q96253	Q96253 plasmodium
3	40	80.0	1516	Q96154	Q96154 plasmodium
4	40	80.0	2206	Q96205	Q96205 plasmodium
5	40	80.0	1182	Q96240	Q96240 plasmodium
6	40	80.0	2340	Q97298	Q97298 plasmodium
7	40	80.0	382	Q04488	Q04488 arabidopsis
8	39	78.0	232	Q76774	Q76774 dictyosteli
9	39	78.0	488	Q98172	Q98172 molluscum c
10	38	76.0	170	Q44172	Q44172 caenorhabdi
11	38	76.0	3973	Q96204	Q96204 plasmodium
12	38	76.0	297	Q97317	Q97317 plasmodium
13	38	76.0	41	Q20155	Q20155 chlorella v
14	37	74.0	1116	Q067838	Q067838 aquifex aeo
15	37	74.0	287	Q93503	Q93503 xenopus lae
16	36	72.0	853	Q206796	Q206796 aquifex aeo
17	36	72.0	154	Q05405	Q05405 anopheles g
18	36	72.0	369	Q21024	Q21024 caenorhabdi
19	36	72.0	1121	Q96192	Q96192 plasmodium
20	36	72.0	2013	Q96216	Q96216 plasmodium
21	36	72.0	1096	Q97257	Q97257 plasmodium
22	36	72.0	644	Q97258	Q97258 plasmodium
23	36	72.0	566	Q92QK5	Q92QK5 arabidopsis
24	35	70.0	270	Q34852	Q34852 bacillus su
25	35	70.0	54	Q07058	Q07058 saccharomyc
26	35	70.0	513	Q15865	Q15865 plasmodium
27	35	70.0	784	Q77339	Q77339 plasmodium
28	35	70.0	467	Q62214	Q62214 caenorhabdi
29	35	70.0	191	Q45682	Q45682 caenorhabdi

30 35 70.0 2539 5 096157 096157 plasmodium  
31 35 70.0 1014 5 096276 096276 plasmodium  
32 35 70.0 1844 5 097287 097287 plasmodium  
33 35 70.0 1542 5 097299 097299 plasmodium  
34 35 70.0 86 8 020119 020119 chlorella v  
35 35 70.0 46 8 020183 020183 chlorella v  
36 35 70.0 100 8 021253 021253 reclinomona  
37 35 70.0 511 8 078330 078330 gonolobus x  
38 35 70.0 85 12 055718 055718 chilo iride  
39 34 68.0 435 2 025942 025942 helicobacte  
40 34 68.0 417 2 051277 051277 borrelia bu  
41 34 68.0 924 2 005842 005842 prevotella  
42 34 68.0 379 2 087181 087181 streptococc  
43 34 68.0 435 2 092379 092379 helicobacte  
44 34 68.0 1407 3 006211 006211 saccharomyc  
45 34 68.0 101 11 008820 008820 mus musculu

ALIGNMENTS

RESULT 1  
Q99248 PRELIMINARY; PRT; 730 AA.  
ID Q99248;  
AC Q99248;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE CHROMOSOME XV READING FRAME ORF YOR019W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DE HAAN M., GRIVELL L.A., MAARSE A.C.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA DE HAAN M., MAARSE A.C., GRIVELL L.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RX MEDLINE; 94019318.  
RA DUMONT M.E., SCHLICHTER J.B., CARDILLO T.S., HAYES M.K., BETHLENDY G.,  
RA SHERMAN F.;  
RL "CYC2 encodes a factor involved in mitochondrial import of yeast  
cytochrome c";  
RL Mol. Cell. Biol. 13:6442-6451(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RX MEDLINE; 94169519.  
RA LEE Y.S., SHIMIZU J., YODA K., YAMASAKI M.;  
RL "Molecular cloning of a gene, DHS1, which complements a  
drug-hypersensitive mutation of the yeast Saccharomyces cerevisiae";  
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).  
DR EMBL; 274927; CAA95209.1; -;  
DR EMBL; X87331; CAA60768.1; -;  
SQ SEQUENCE 730 AA; 83365 MW; 09D72CD9 CRC32;

Query Match 84.0%; Score 42; DB 3; Length 730;  
Best Local Similarity 90.0%; Pred. No. 21;  
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 KKKKKKFLFL 10  
||||| |||

Db 717 KKKKKSLFL 726

## RESULT 2

O96253 ID 096253 PRELIMINARY; PRT; 1817 AA.  
AC 096253;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 220.3 KD PROTEIN.  
GN PFB0800C.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum";  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001418; AAC71949.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 1817 AA; 220263 MW; 53752050 CRC32;

Query Match 82.0%; Score 41; DB 5; Length 1817;

Best Local Similarity 70.0%; Pred. No. 62;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10

|||||

Db 666 KKKKKFFYI 675

## RESULT 3

O96154 ID 096154 PRELIMINARY; PRT; 1516 AA.  
AC 096154;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE RAD2 ENDONUCLEASE.  
GN PFB0265C.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum";  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001383; AAC71842.1; -.  
KW Endonuclease.  
SQ SEQUENCE 1516 AA; 178730 MW; E899951C CRC32;

Query Match 80.0%; Score 40; DB 5; Length 1516;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFL 8

|||||

Db 700 KKKKKFL 707

## RESULT 4

O96205 ID 096205 PRELIMINARY; PRT; 2206 AA.  
AC 096205;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 264.1 KD PROTEIN.  
GN PFB0560W.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum";  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001403; AAC71901.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 2206 AA; 264100 MW; AEE134B0 CRC32;

Query Match 80.0%; Score 40; DB 5; Length 2206;

Best Local Similarity 60.0%; Pred. No. 1e+02;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10

|||||

Db 2189 KKKKKYIYI 2198

## RESULT 5

O96240 ID 096240 PRELIMINARY; PRT; 1182 AA.  
AC 096240;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE PREDICTED INTEGRAL MEMBRANE PROTEIN.  
GN PFB0735C.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum";  
RL Science 282:1126-1132(1998).  
DR EMBL: AE001414; AAC71936.1; -.  
SQ SEQUENCE 1182 AA; 144123 MW; 98F97AF0 CRC32;

Query Match 80.0%; Score 40; DB 5; Length 1182;

Best Local Similarity 60.0%; Pred. No. 62;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10

|||||

Db 657 KKKKKYIYI 666

## RESULT 6

O97298 ID 097298 PRELIMINARY; PRT; 2340 AA.  
AC 097298;

DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE MAL3P7.34 PROTEIN.  
 GN MAL3P7.34.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3D7;  
 RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL034559; CAB39045.1; -;  
 SQ SEQUENCE 2340 AA; 280914 MW; 44043283 CRC32;

Query Match 80.0%; Score 40; DB 5; Length 2340;  
 Best Local Similarity 88.9%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 ||||| ||  
 Db 995 KKKKKKILF 1003

RESULT 7  
 O04488  
 ID 004488 PRELIMINARY; PRT; 382 AA.  
 AC 004488;  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)  
 DE SEQUENCE OF BAC F21M12 FROM ARABIDOPSIS THALIANA CHROMOSOME 1,  
 DE COMPLETE SEQUENCE.  
 GN F21M12.4.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA VYSOTSKAIA V.S., OSBORNE B.I., TORIUMI M., YU G., QJI O., SHEN Y.K.,  
 RA ARAUJO R., AU M., RUEHLER E., CONWAY A.B., CONWAY A.R., DEWAR K.,  
 RA FENG J., KIM C., KURTZ D., LI Y., SHINN P., SUN H., DAVIS R.W.,  
 RA ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC000132; AAB60722.1; -;  
 SQ SEQUENCE 382 AA; 43555 MW; D732D7FB CRC32;

Query Match 80.0%; Score 40; DB 10; Length 382;  
 Best Local Similarity 88.9%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 ||||| ||  
 Db 332 KKKKKKLLF 340

RESULT 8  
 O76774  
 ID 076774 PRELIMINARY; PRT; 232 AA.  
 AC 076774;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE GTP BINDING PROTEIN RARE7L.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Dictyostellida; Dictyostelium.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=AX2;  
 RA CLOUGHERTY C., GERISCH G.;  
 RT "Characterization of a Rab-like GTPase from Dictyostelium  
 RT discoideum";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF081681; AAC34837.1; -;  
 DR PFAM; PF00071; ras; 1.  
 SQ SEQUENCE 232 AA; 27123 MW; 47278584 CRC32;

Query Match 78.0%; Score 39; DB 5; Length 232;  
 Best Local Similarity 70.0%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 10  
 ||||| ||  
 Db 213 KKKKKKYLII 222

RESULT 9  
 Q98172  
 ID 098172 PRELIMINARY; PRT; 488 AA.  
 AC 098172;  
 DT 01-FEB-1997 (TReMBLrel. 02, Created)  
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MC001R.  
 GN MC001R.  
 OS Molluscum contagiosum virus subtype 1 (MCV1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Molluscipoxvirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE; 96325459.  
 RA SENKEVICH T.G., BUGERT J.J., SISLER J.R., KOONIN E.V., DARAI G.,  
 RA MOSS B.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U60315; AAC55129.1; -;  
 SQ SEQUENCE 488 AA; 54066 MW; B90B4E9C CRC32;

Query Match 78.0%; Score 39; DB 12; Length 488;  
 Best Local Similarity 80.0%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 10  
 ||||| ||  
 Db 20 KKKKKGFFFL 29

RESULT 10  
 O44172  
 ID 044172 PRELIMINARY; PRT; 170 AA.  
 AC 044172;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE F58F6.3 PROTEIN.  
 GN F58F6.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Peloderfinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 94150718.

Wed Sep 8 09:45:16 1999

WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SALLDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RL elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA DANTE M., KRAMER J., GIBSON A.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF036699; AAB88357.1; - 5F4A7EC8 CRC32;  
 SQ SEQUENCE 170 AA; 19554 MW; 5F4A7EC8 CRC32;

Query Match 76.0%; Score 38; DB 5; Length 170;  
 Best Local Similarity 80.0%; Pred. NO. 26;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10  
 ||||| ||  
 DB 81 KKKKKKLSFL 90

RESULT 11  
 O96204 PRELIMINARY; PRT; 3973 AA.  
 ID O96204;  
 AC O96204;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE HYPOTHETICAL 480.3 KD PROTEIN.  
 GN PFB0555C.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,  
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEDERSON J.,  
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,  
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE O., SMITH H.O.,  
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;  
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium  
 RL Science 282:1126-1132(1998).  
 RL EMBL: AE001402; AAC71900.1; -  
 DR Hypothetical protein.  
 KW SEQUENCE 3973 AA; 480283 MW; C850AE7A CRC32;  
 SQ SEQUENCE 3973 AA; 480283 MW; C850AE7A CRC32;

Query Match 76.0%; Score 38; DB 5; Length 3973;  
 Best Local Similarity 60.0%; Pred. NO. 3.4e+02;  
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10  
 ||:|:|:|  
 DB 2630 KKKKKYIFL 2639

RESULT 12  
 O97317

ID O97317 PRELIMINARY; PRT; 297 AA.  
 AC O97317;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE MAL3P7.18 PROTEIN.  
 DT MAL3P7.18  
 GN Plasmodium falciparum.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-3D7;  
 RA OLIVER K., BOWMAN S., HARRIS D., LAWSON D., QUAIL M., BARRELL B.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL034559; CAB39066.1; - 7FC2C40B CRC32;  
 SQ SEQUENCE 297 AA; 35625 MW; 7FC2C40B CRC32;

Query Match 76.0%; Score 38; DB 5; Length 297;  
 Best Local Similarity 66.7%; Pred. NO. 41;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 |||||:|:  
 DB 289 KKKKKKIYI 297

RESULT 13  
 O20155 PRELIMINARY; PRT; 41 AA.  
 ID O20155;  
 AC O20155;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE ORF41C.  
 OS Chlorella vulgaris.  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE: 97303241.  
 RA WAKASUGI T., NAGAI T., KAPOOR M., SUGITA M., ITO M., ITO S.,  
 RA TSUDZUKI J., NAKASHIMA K., TSUDZUKI T., SUZUKI Y., HAMADA A., OHYA T.,  
 RA INAMURA A., YOSHINAGA K., SUGIURA M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 DR EMBL: AB001684; BAR20664.1; -  
 KW Chloroplast.  
 SQ SEQUENCE 41 AA; 5328 MW; 31B8B267 CRC32;

Query Match 76.0%; Score 38; DB 8; Length 41;  
 Best Local Similarity 71.4%; Pred. NO. 8.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 KKKKKK----FLFL 10  
 ||||| ||||  
 DB 17 KKKKKKNNVEFLFL 30

RESULT 14  
 O67838 PRELIMINARY; PRT; 1116 AA.  
 ID O67838;  
 AC O67838;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 127.8 KD PROTEIN.  
 GN AQ\_2054.  
 OS Aquifex aeolicus.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RX MEDLINE; 9819666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL Nature 392:353-358(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AE000770; AAC07805.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 1116 AA; 127794 MW; 04851157 CRC32;

Query Match 74.08; Score 37; DB 2; Length 1116;  
 Best Local Similarity 87.5%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFL 8  
 :|||||  
 Db 84 RKKKKKFL 91

RESULT 15  
 O93503  
 ID O93503 PRELIMINARY; PRT; 287 AA.  
 AC O93503;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE MYRISTOYLATED ALANINE-RICH C KINASE SUBSTRATE.  
 GN MARCKS.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;  
 OC Xenopus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98030614.  
 RA SHI Y., SULLIVAN S.K., PITTERLE D.M., KENNINGTON E.A., GRAFF J.M.,  
 RA BLACKSHEAR P.J.;  
 RT "Mechanisms of MARCKS gene activation during Xenopus development.";  
 RL J. Biol. Chem. 272:29390-29390(1997).  
 DR EMBL; AF017299; AAC61897.1; -;  
 DR PROSITE; PS00826; MARCKS\_1; 1.  
 DR PROSITE; PS00827; MARCKS\_2; 1.  
 SQ SEQUENCE 287 AA; 29147 MW; 50689EEE CRC32;

Query Match 74.08; Score 37; DB 13; Length 287;  
 Best Local Similarity 77.8%; Pred. No. 58;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 :|||||  
 Db 125 KKKKKRFSF 133

Search completed: September 7, 1999, 23:17:32  
 Job time: 1775 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5

Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 1999, 00:42:59 ; Search time 21.13 Seconds  
(without alignments)  
13.378 Million cell updates/sec

Title: US-09-124-280A-42

Perfect score: 50

Sequence: 1 KKKKKKFLFL 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	74.0	331	1 MACS_BOVIN	P12624 bos taurus
2	37	74.0	280	1 MACS_CHICK	P16527 gallus gall
3	37	74.0	331	1 MACS_HUMAN	P29966 homo sapien
4	37	74.0	308	1 MACS_MOUSE	P26645 mus musculu
5	37	74.0	308	1 MACS_RAT	P30009 rattus norv
6	36	72.0	830	1 PAML_YEAST	P37304 saccharomyc
7	36	72.0	2875	1 RRPL_TSWV1	P28976 tomato spot
8	35	70.0	194	1 MRP_HUMAN	P49006 homo sapien
9	35	70.0	199	1 MRP_MOUSE	P28667 mus musculu
10	35	70.0	198	1 MRP_RABIT	P35566 oryctolagus
11	35	70.0	1386	1 RPOD_MARPO	P06274 marchantia
12	35	70.0	948	1 RPOD_PODAN	Q01521 podospora a
13	35	70.0	233	1 RR2_CYACA	P35014 cyanidium c
14	35	70.0	240	1 Y482_METJA	Q57905 methanococc
15	35	70.0	2136	1 YCF2_MARPO	P09975 marchantia
16	34	68.0	326	1 YA55_METJA	Q58455 methanococc
17	34	68.0	766	1 YS83_CAEEL	Q09622 caenorhabdi
18	33	66.0	3229	1 BRG2_MOUSE	P97929 mus musculu
19	33	66.0	489	1 COAT_FWVD	P09519 figwort mos
20	33	66.0	268	1 ESL2_MYCGE	Q49418 mycoplasma
21	33	66.0	312	1 IF2B_DROME	P41375 drosophila
22	33	66.0	539	1 PNUT_DROME	P40797 drosophila
23	33	66.0	606	1 RA17_SCHPO	P50531 schizosacch
24	33	66.0	175	1 Y433_METJA	Q57875 methanococc
25	33	66.0	197	1 YMBV_YEAST	Q03559 saccharomyc
26	32	64.0	333	1 IF2B_HUMAN	P20042 homo sapien
27	32	64.0	333	1 IF2B_RABIT	P41035 oryctolagus
28	32	64.0	270	1 IF2B_WHEAT	Q24473 triticum ae
29	32	64.0	229	1 REPX_STAAU	P03862 staphylococ
30	32	64.0	2339	1 RPCI_PLAFA	P27625 plasmodium
31	32	64.0	815	1 RGR_HUMAN	Q92834 homo sapien
32	32	64.0	256	1 TAGA_BACSU	P27620 bacillus su
33	32	64.0	114	1 VG40_BPT4	P17171 bacterioph
34	32	64.0	1050	1 Y032_HUMAN	Q15034 homo sapien
35	32	64.0	233	1 YEG7_YEAST	P39983 saccharomyc
36	32	64.0	167	1 YK20_YEAST	P36133 saccharomyc
37	32	64.0	1545	1 YMH2_YEAST	Q03640 saccharomyc
38	32	64.0	435	1 YMJ8_CAEEL	P34483 caenorhabdi
39	31	62.0	737	1 ADDA_HUMAN	P35611 homo sapien
40	31	62.0	726	1 ADDB_HUMAN	P35612 homo sapien
41	31	62.0	168	1 CBP2_DICDI	P54653 dictyostell
42	31	62.0	2039	1 CCHI_YEAST	P50077 saccharomyc
43	31	62.0	488	1 COAT_CAMVD	P03543 cauliflower

44 31 62.0 490 1 COAT\_CAMVD P03544 cauliflower  
45 31 62.0 574 1 VGLF\_HRSVR P11209 human respi

## ALIGNMENTS

RESULT 1  
ID MACS\_BOVIN STANDARD; PRT; 331 AA.  
AC P12624;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS) (ACAMP-81).  
GN MACS.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89282412.  
RA STUMPO D.J., GRAFF J.M., ALBERT K.A., GRENGARD P., BLACKSHEAR P.J.;  
RT "Nucleotide sequence of a cDNA for the bovine myristoylated  
alanine-rich C kinase substrate (MARCKS).";  
RL NUCLEIC ACIDS RES. 17:3987-3988(1989).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 89264553.  
RA STUMPO D.J., GRAFF J.M., ALBERT K.A., GRENGARD P., BLACKSHEAR P.J.;  
RT "Molecular cloning, characterization, and expression of a cDNA  
encoding the '80- to 87-kDa' myristoylated alanine-rich C kinase  
substrate: a major cellular substrate for protein kinase C.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:4012-4016(1989).  
RN [3]  
RP PARTIAL SEQUENCE.  
RX MEDLINE; 92171958.  
RA MIZUTANI A., TOKUMITSU H., HIDAKA H.;  
RT "Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein  
interacting with synapsin I";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 182:1395-1401(1992).  
RN [4]  
RP PHOSPHORYLATION SITES.  
RX MEDLINE; 89308594.  
RA GRAFF J.M., STUMPO D.J., BLACKSHEAR P.J.;  
RT "Characterization of the phosphorylation sites in the chicken and  
bovine myristoylated alanine-rich C kinase substrate protein, a  
prominent cellular substrate for protein kinase C.";  
RL J. BIOL. CHEM. 264:11912-11919(1989).  
RN [5]  
RP PHOSPHORYLATION SITES, AND REVISIONS.  
RX MEDLINE; 94308052.  
RA TANIGUCHI H., MANENTI S., SUZUKI M., TITANI K.;  
RT "Myristoylated alanine-rich C kinase substrate (MARCKS), a major  
protein kinase C substrate, is an in vivo substrate of  
proline-directed protein kinase(s). A mass spectroscopic analysis of  
the post-translational modifications.";  
RL J. BIOL. CHEM. 269:18299-18302(1994).  
RN [6]  
RP REVERSIBLE ASSOCIATION WITH THE MEMBRANE.  
RX MEDLINE; 91238951.  
RA THELEN M., ROSEN A., NAIRN A.C., ADEREM A.;  
RT "Regulation by phosphorylation of reversible association of a  
myristoylated protein kinase C substrate with the plasma membrane.";  
RL NATURE 351:320-322(1991).  
RN [7]  
RP ACTIN-FILAMENT CROSS-LINKING.  
RX MEDLINE; 92220195.  
RA HARTWIG J.H., THELEN M., ROSEN A., JANNEY P.A., NAIRN A.C.,  
ADEREM A.;  
RT "MARCKS is an actin filament crosslinking protein regulated by  
protein kinase C and calcium-calmodulin.";

CC CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
CC CC -1- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC CC  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC CC the European Bioinformatics Institute. There are no restrictions on its  
CC CC use by non-profit institutions as long as its content is in no way  
CC CC modified and this statement is not removed. Usage by and for commercial  
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC CC  
CC CC EMBL; M31650; G212288; -  
CC CC PIR; A41400; A41400.  
CC CC DR PROSITE; PS00826; MARCKS\_1; 1.  
CC CC DR PROSITE; PS00827; MARCKS\_2; 1.  
CC CC DR HSP; P04002; IYFA.  
CC CC KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
CC CC MEMBRANE.  
CC CC FT INIT\_MET 0 0  
CC CC FT LIPID 1 1 MYRISTATE.  
CC CC FT DOMAIN 116 140 CALMODULIN-BINDING (PSD).  
CC CC FT MOD\_RES 123 123 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 127 127 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 131 131 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 134 134 PHOSPHORYLATION (BY PKC).  
CC CC SQ SEQUENCE 280 AA; 27597 MW; 41D34538 CRC32;  
  
Query Match 74.0%; Score 37; DB 1; Length 280;  
Best Local Similarity 77.8%; Pred. No. 9.6;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KKKKKKFLF 9  
DB 116 KKKKKRFSF 124  
  
RESULT 3  
MACS\_HUMAN STANDARD; PRT; 331 AA.  
AC P29986;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)  
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS) (PROTEIN KINASE  
C SUBSTRATE, 80 KD PROTEIN, LIGHT CHAIN) (PKCSL) (80K-L PROTEIN).  
GN MACS.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91317795.  
RA HARLAN D.M., GRAFF J.M., STUMPO D.J., EDDY R.L. JR., SHOWS T.B.,  
RA BOYLE J.M., BLACKSHEAR P.J.;  
RT "The human myristoylated alanine-rich C kinase substrate (MARCKS)  
gene (MACS). Analysis of its gene product, promoter, and chromosomal  
localization.";  
RT J. BIOL. CHEM. 266:14399-14405(1991).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 93052291.  
RX SAKAI K., HIRAI M., KUDOH J., MINOSHIMA S., SHIMIZU N.;  
RA "Molecular cloning and chromosomal mapping of a cDNA encoding human  
80K-L protein: major substrate for protein kinase C.";  
RT GENOMICS 14:175-178(1992).  
RL  
CC -1- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
CC CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
CC CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
CC CC -1- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
CC CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
CC CC -1- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC CC

RL NATURE 356:618-622(1992).  
CC CC -1- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
CC CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
CC CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
CC CC -1- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
CC CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
CC CC -1- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC CC -1- CAUTION: REF.1 AND REF.2 SEQUENCES DIFFER FROM THAT SHOWN IN  
CC CC POSITIONS 140 TO 150 DUE TO A FRAMESHIFT.  
CC CC  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC CC the European Bioinformatics Institute. There are no restrictions on its  
CC CC use by non-profit institutions as long as its content is in no way  
CC CC modified and this statement is not removed. Usage by and for commercial  
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC CC  
CC CC EMBL; M24638; G163340; ALT\_FRAME.  
CC CC PIR; A32904; A32904.  
CC CC DR PIR; PS0338; PS0338.  
CC CC DR PIR; S08341; S08341.  
CC CC DR PROSITE; PS00826; MARCKS\_1; 1.  
CC CC DR PROSITE; PS00827; MARCKS\_2; 1.  
CC CC KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
CC CC MEMBRANE.  
CC CC FT INIT\_MET 0 0  
CC CC FT LIPID 1 1 MYRISTATE.  
CC CC FT DOMAIN 150 174 CALMODULIN-BINDING (PSD).  
CC CC FT MOD\_RES 26 26 PHOSPHORYLATION.  
CC CC FT MOD\_RES 45 45 PHOSPHORYLATION.  
CC CC FT MOD\_RES 80 80 PHOSPHORYLATION.  
CC CC FT MOD\_RES 99 99 PHOSPHORYLATION.  
CC CC FT MOD\_RES 116 116 PHOSPHORYLATION.  
CC CC FT MOD\_RES 133 133 PHOSPHORYLATION.  
CC CC FT MOD\_RES 157 157 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 161 161 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 165 165 PHOSPHORYLATION (BY PKC).  
CC CC FT MOD\_RES 168 168 PHOSPHORYLATION (BY PKC).  
CC CC SQ SEQUENCE 331 AA; 31450 MW; 27E69E59 CRC32;  
  
Query Match 74.0%; Score 37; DB 1; Length 331;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 KKKKKKFLF 9  
DB 150 KKKKKRFSF 158  
  
RESULT 2  
MACS\_CHICK STANDARD; PRT; 280 AA.  
ID MACS\_CHICK  
AC P16527;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).  
OS GALLUS GALLUS (CHICKEN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 90114197.  
RA GRAFF J.M., STUMPO D.J., BLACKSHEAR P.J.;  
RT "Molecular cloning, sequence, and expression of a cDNA encoding the  
chicken myristoylated alanine-rich C kinase substrate (MARCKS).";  
RL MOL. ENDOCRINOL. 3:1903-1906(1989).  
CC -1- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
CC CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
CC CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
CC CC -1- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
CC CC



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: M68956; G187387; -;  
 DR EMBL: M68955; G187385; -;  
 DR EMBL: D10522; G219894; -;  
 DR PIR: A38873; A38873.  
 DR MM: 177061; -;  
 DR PROSITE; PS00826; MARCKS\_1; 1.  
 DR PROSITE; PS00827; MARCKS\_2; 1.  
 KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
 KW MEMBRANE.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 CALMODULIN-BINDING (PSD).  
 FT DOMAIN 151 175 CALMODULIN-BINDING (PSD).  
 FT MOD\_RES 158 158 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 166 166 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 169 169 PHOSPHORYLATION (BY PKC).  
 FT MOD\_RES 169 169 PHOSPHORYLATION (BY PKC).  
 FT CONFLICT 83 83 S -> A (IN REF. 1).  
 FT CONFLICT 118 118 A -> P (IN REF. 1).  
 FT CONFLICT 233 233 P -> S (IN REF. 1).  
 FT CONFLICT 286 307 PGAPPEGEAAPAEPAARASS ->  
 FT SEQUENCE 331 AA; 31413 MW; 152CCCA2 CRC32;  
 LCPRGSGSPRGAGRRSLNQ (IN REF. 1).

Query Match 74.0% Score 37; DB 1; Length 331;  
 Best Local Similarity 77.8%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
 |||||  
 Db 151 KKKKKRFSF 159

RESULT 4.  
 MACS\_MOUSE  
 ID MACS\_MOUSE STANDARD; PRT; 308 AA.  
 AC P26645;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DE 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)  
 DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).  
 GN MACS.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-MACROPHAGE;  
 RX MEDLINE: 91172836.  
 RA SEYKORA J.T., RAVETCH J.V., ADEREM A.;  
 RT "Cloning and molecular characterization of the murine macrophage '68-  
 RT kda' protein kinase C substrate and its regulation by bacterial  
 RT lipopolysaccharide." RT  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:2505-2509(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-FIBROBLAST;  
 RX MEDLINE: 91330872.  
 RA BROOKS S.F., HERGET T., ERUSALIMSKY J.D., ROZENGURT E.;  
 RT "Protein kinase C activation potentially down-regulates the expression  
 RT of its major substrate, 80K, in Swiss 3T3 cells.";  
 RL EMBO J. 10:2497-2505(1991).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC STRAIN-SWISS; TISSUE-FIBROBLAST;

RX MEDLINE: 90346162.  
 RA BROOKS S.F., ERUSALIMSKY J.D., TOTTY N.F., ROZENGURT E.;  
 RT "Purification and internal amino acid sequence of the 80 kda protein  
 RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with  
 RT substrates from brain.";  
 RL FEBS LETT. 268:291-293(1990).  
 CC -1- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
 CC -1- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND  
 CC VERY LOW LEVELS IN LIVER.  
 CC -1- INDUCTION: BY LIPOPOLYSACCHARIDE.  
 CC -1- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M60474; G199027; -;  
 CC PIR: A39169; A39169.  
 CC PIR: S16519; S16519.  
 CC MGD: MGI:96907; MACS.  
 CC PROSITE; PS00826; MARCKS\_1; 1.  
 CC PROSITE; PS00827; MARCKS\_2; 1.  
 KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;  
 KW MEMBRANE.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT DOMAIN 144 168 CALMODULIN-BINDING (PSD).  
 FT MOD\_RES 151 151 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 155 155 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 159 159 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT CONFLICT 95 97 AGA -> TGT (IN REF. 2).  
 FT SEQUENCE 308 AA; 29530 MW; A91BD349 CRC32;

Query Match 74.0% Score 37; DB 1; Length 308;  
 Best Local Similarity 77.8%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
 |||||  
 Db 144 KKKKKRFSF 152

RESULT 5  
 MACS\_RAT  
 ID MACS\_RAT STANDARD; PRT; 308 AA.  
 AC P30009;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
 DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).  
 GN MACS.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE: 91201362.  
 RA ERUSALIMSKY J.D., BROOKS S.F., HERGET T., MORRIS C., ROZENGURT E.;  
 RT "Molecular cloning and characterization of the acidic 80-kda protein  
 RT kinase C substrate from rat brain. Identification as a  
 RT glycoprotein.";  
 RL J. BIOL. CHEM. 266:7073-7080(1991).

CC 2A DEPLETION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X73454; G456667; -

CC EMBL; Z70202; E228706; -

CC EMBL; Z68329; E215297; -

CC PIR; S45179; S45179; -

CC SGD; L0001334; PAM1.

CC -----

CC KW COILED COIL. 379 400 COILED COIL.

CC FT DOMAIN 481 514 COILED COIL.

CC FT DOMAIN 515 522 POLY-GLN.

CC FT DOMAIN 810 830 ARG/LYS-RICH (HIGHLY BASIC).

CC FT SEQUENCE 830 AA; 92886 MW; D08F1CD9 CRC32;

CC SQ

Query Match 72.0%; Score 36; DB 1; Length 830;

Best Local Similarity 77.8%; Pred. No. 41;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9

DB 816 EKKKKKFSF 824

RESULT 7

RRPL\_TSWV1 STANDARD; PRT; 2875 AA.

ID RRPL\_TSWV1

AC P28976;

DT 01-DEC-1992 (REL. 24, CREATED)

DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (L PROTEIN).

GN L.

OS TOMATO SPOTTED WILT VIRUS (STRAIN BRAZILIAN BR-01) (TSWV).

OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; BUNYAVIRIDAE; TOSPOVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91374019.

RA DE HAAN P., KORMELINK R., DE OLIVEIRA RESENDE R., VAN POELWIJK F.,

RA PETERS D., GOLDBACH R.;

RT "Tomato spotted wilt virus L RNA encodes a putative RNA polymerase."

RL J. GEN. VIROL. 72:2207-2216(1991).

CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE

CC + RNA(N).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; D10066; G222681; -

CC PIR; JQ1335; RRVUTW.

CC KW TRANSFERASE; NUCLEOTIDYLTRANSFERASE; RNA-DIRECTED RNA POLYMERASE.

CC SQ SEQUENCE 2875 AA; 331498 MW; 3B0A92AE CRC32;

Query Match 72.0%; Score 36; DB 1; Length 2875;

Best Local Similarity 77.8%; Pred. No. 1.4e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9

DB 144 KKKKKKFSF 152

RN PHOSPHORYLATION SITES.

RC TISSUE-BRAIN;

RX MEDLINE; 93135774.

RA HEEMSKERK F.M., CHEN H.C., HUANG F.L.;

RT Ser-160 of MARKS C phosphorylates Ser-152, Ser-156 and Ser-163 but not

RL BIOCHEM. BIOPHYS. RES. COMMUN. 190:236-241(1993).

CC -!- FUNCTION: MARKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR

CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND

CC SYNAPSIN. MARKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.

CC -!- PIM: PHOSPHORYLATION BY PKC DISPLACES MARKS FROM THE MEMBRANE. IT

CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.

CC -!- SIMILARITY: BELONGS TO THE MARKS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M59859; -; NOT\_ANNOTATED\_CDS.

CC PIR; A39773; A39773.

CC PROSITE; PS00826; MARKS\_1; 1.

CC PROSITE; PS00827; MARKS\_2; 1.

CC KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING; ACTIN-BINDING;

CC MEMBRANE.

CC FT INIT\_MET 0 0 BY SIMILARITY.

CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).

CC FT DOMAIN 144 168 CALMODULIN-BINDING (PSD).

CC FT MOD\_RES 151 151 PHOSPHORYLATION (BY PKC).

CC FT MOD\_RES 155 155 PHOSPHORYLATION (BY PKC).

CC FT MOD\_RES 162 162 PHOSPHORYLATION (BY PKC).

CC FT SEQUENCE 308 AA; 29663 MW; A6B1CF2A CRC32;

Query Match 74.0%; Score 37; DB 1; Length 308;

Best Local Similarity 77.8%; Pred. No. 11;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9

DB 144 KKKKKKFSF 152

RESULT 6

PAM1\_YEAST STANDARD; PRT; 830 AA.

ID PAM1\_YEAST

AC P37304;

DT 01-OCT-1994 (REL. 30, CREATED)

DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE PAM1 PROTEIN.

GN PAM1 OR YDR251W OR YD8419.18 OR YD9320A.01.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;

OC SACCHAROMYCETACEAE; SACCHAROMYCES.

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN-W303-1A;

RA MEDLINE; 94148839.

RA HU G.-Z., RONNE H.;

RT "Overexpression of yeast PAM1 gene permits survival without protein

RL phosphatase 2A and induces a filamentous phenotype."

RL J. BIOL. CHEM. 269:3429-3435(1994).

CC -!- FUNCTION: NOT KNOWN. IT IS A SUPPRESSOR OF PROTEIN PHOSPHATASE

```

Db 2669 KRKKKGFLF 2677

RESULT 8
MRP_HUMAN
ID MRP_HUMAN STANDARD; PRT; 194 AA.
AC P49006;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MARCKS-RELATED PROTEIN (MAC-MARCKS).
GN MLP OR MRP.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP BLOCKX H., MAERTENS C., FRANSEN L.M.L.;
RL SUBMITTED (FEB-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND
CC CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.
CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X70326; G38435; -
CC MIM; 602940; -
CC DR PROSITE; PS00826; MARCKS_1; 1.
CC DR PROSITE; PS00827; MARCKS_2; 1.
CC KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT DOMAIN 86 99 CALMODULIN-BINDING (PSD).
CC FT MOD_RES 92 92 POLY-LYS.
CC FT MOD_RES 92 92 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
CC FT MOD_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
CC FT MOD_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
CC SQ SEQUENCE 194 AA; 19397 MW; DAD367FF CRC32;

Query Match 70.08; Score 35; DB 1; Length 194;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KRKKKGFLF 9
| | | | |
Db 86 KRKKKFSF 93

RESULT 9
MRP_MOUSE
ID MRP_MOUSE STANDARD; PRT; 199 AA.
AC P28667;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE MARCKS-RELATED PROTEIN (MAC-MARCKS) (BRAIN PROTEIN F52).
GN MLP OR MRP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=BRAIN;
RX MEDLINE; 91323504.
RA UMEKAGE T., KATO K.;

```

"A mouse brain cDNA encodes a novel protein with the protein kinase C phosphorylation site domain common to MARCKS.";

FEBS LETT. 286:147-151(1991).

[2]

SEQUENCE FROM N.A.

RP MEDLINE; 92386598.

RA LI J., ADEREM A.;

RT "MacMARCKS, a novel member of the MARCKS family of protein kinase C substrates.";

RL CELL 70:791-801(1992).

RN [3]

SEQUENCE FROM N.A.

RP MEDLINE; 94010885.

RA LOBACH D.F., ROCHELLE J.M., WATSON M.L., SELDIN M.F., BLACKSHEAR P.J.;

RT "Nucleotide sequence, expression, and chromosomal mapping of Mrp and mapping of five related sequences.";

RL GENOMICS 17:194-204(1993).

RN [4]

CHARACTERIZATION.

RP MEDLINE; 92317080.

RA BLACKSHEAR P.J., VERGHESE G.M., JOHNSON J.D., HAUPT D.M., STUMPO D.J.;

RT "Characteristics of the P52 protein, a MARCKS homologue.";

RL J. BIOL. CHEM. 267:13540-13546(1992).

CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.

CC -!- TISSUE SPECIFICITY: BRAIN (MOSTLY IN DENTATE GYRUS, ANTERIOR OLFACTORY NUCLEUS, PRIMARY OLFACTORY CORTEX, ENTORHINAL CORTEX, MEDIAL PREOPTIC AREA, AND DORSOMEDIAL HYPOTHALAMIC NUCLEUS).

CC -!- PTM: PHOSPHORYLATED.

CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

-----

EMBL; X61399; G50944; -

DR EMBL; S65597; E88783; -

DR PIR; S17185; S17185.

DR PIR; B43341; B43341.

DR MGD; MGI-97143; MLP.

DR PROSITE; PS00826; MARCKS\_1; 1.

DR PROSITE; PS00827; MARCKS\_2; 1.

KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.

FT INIT\_MET 0 0 BY SIMILARITY.

FT LIPID 1 1 MYRISTATE.

FT DOMAIN 86 99 CALMODULIN-BINDING (PSD).

FT MOD\_RES 92 92 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

FT MOD\_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

FT MOD\_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

SQ SEQUENCE 199 AA; 20034 MW; 5CC753C4 CRC32;

Query Match 70.08; Score 35; DB 1; Length 199;

Best Local Similarity 87.5%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KRKKKGFLF 9

| | | | |

Db 86 KRKKKFSF 93

RESULT 10

MRP\_RABIT

ID MRP\_RABIT STANDARD; PRT; 198 AA.

AC P35566;

DT 01-JUN-1994 (REL. 29, CREATED)

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE MARCKS-RELATED PROTEIN (MAC-MARCKS).  
GN MLP OR MRP.  
OS ORYCTOLAGUS CUNICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX TISSUE-MACROPHAGE;  
RC MEDLINE; 92386598.  
RA LI J., ADEREM A.;  
RT "MACMARCKS, a novel member of the MARCKS family of protein kinase C  
RT substrates".  
RL CELL 791-801(1992).  
CC -!- FUNCTION: MAY BE INVOLVED IN COUPLING THE PROTEIN KINASE C AND  
CC CALMODULIN SIGNAL TRANSDUCTION SYSTEMS.  
CC -!- PTM: PHOSPHORYLATED.  
CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; S43921; G255036; -.  
DR PIR; A43341; A43341.  
DR PROSITE; PS00826; MARCKS\_1; 1.  
DR PROSITE; PS00827; MARCKS\_2; 1.  
KW PHOSPHORYLATION; MYRISTYLATION; CALMODULIN-BINDING.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
FT DOMAIN 86 99 CALMODULIN-BINDING (FSD).  
FT MOD\_RES 92 92 PHOSPHORYLATION (BY PKC). (BY SIMILARITY).  
FT MOD\_RES 100 100 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
FT MOD\_RES 103 103 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
SQ SEQUENCE 198 AA; 19635 MW; 04B10545 CRC32;  
  
Query Match 70.0%; Score 35; DB 1; Length 198;  
Best Local Similarity 87.5%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 KKKKKFLF 9  
Db 86 KKKKKFSF 93  
  
RESULT 11  
RPOD\_MARPO STANDARD; PRT; 1386 AA.  
AC F06274;  
DT 01-JAN-1988 (REL. 06, CREATED)  
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)  
DE DNA-DIRECTED RNA POLYMERASE BETA" CHAIN (EC 2.7.7.6).  
GN RPOC2.  
OS MARCHANTIA POLYMORPHA (LIVERWORT).  
OC CHLOROPLAST.  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; MARCHANTIOPSIDA;  
OC MARCHANTIALES; MARCHANTIACEAE; MARCHANTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA OHYAMA K.;  
RL SUBMITTED (OCT-1986) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP COMPLETE GENOME.  
RA OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SANO T., SANO S.,  
RA UMESONO K., SHIKI Y., TAKEUCHI M., CHANG Z., AOTA S., INOKUCHI H.,  
RA OZEKI H.;  
RT "Chloroplast gene organization deduced from complete sequence of  
RT liverwort Marchantia polymorpha chloroplast DNA.";

RL NATURE 322:572-574(1986).  
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES.  
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +  
CC RNA(N).  
CC -!- SUBUNIT: IN CHLOROPLAST THE RNA POLYMERASE IS COMPOSED OF FOUR  
CC SUBUNITS: ALPHA, BETA, BETA', AND BETA".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X04465; G11650; -.  
DR PIR; A00698; RNLVC2.  
DR PIR; S01575; S01575.  
DR PFAM; PF00623; RNA\_pol\_A; 1.  
DR MENDEL; 4115; MARPOI\_POC2; 1.  
KW TRANSCRIPTION; DNA-DIRECTED RNA POLYMERASE; CHLOROPLAST.  
SQ SEQUENCE 1386 AA; 160155 MW; 203F2178 CRC32;  
  
Query Match 70.0%; Score 35; DB 1; Length 1386;  
Best Local Similarity 77.8%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 KKKKKFLFL 10  
Db 562 KKKKNFYFL 570  
  
RESULT 12  
RPOD\_PODAN STANDARD; PRT; 948 AA.  
AC Q01521;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PROBABLE DNA-DIRECTED RNA POLYMERASE (EC 2.7.7.6).  
OS PODOSPORA ANSERINA.  
OG MITOCHONDRION.  
OC PLASMOD PAL2-1.  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;  
OC SORDARIALES; SORDARIACEAE; PODOSPORA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AL2;  
RX MEDLINE; 93113721.  
RA HERMANS J., OSTEWACZ H.D.;  
RT "The linear mitochondrial plasmid PAL2-1 of a long-lived Podospora  
RT anserina mutant is an inversion encoding a DNA and RNA polymerase.";  
RL CURR. GENET. 22:491-500(1992).  
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES.  
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +  
CC RNA(N).  
CC -!- SIMILARITY: BELONGS TO THE PHAGE AND MITOCHONDRIAL RNA POLYMERASES  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X60707; E38184; -.

```

DR PROSITE; PS00489; RNA_POL_PHAGE_2; 1.
DR PROSITE; PS00900; RNA_POL_PHAGE_1; 1.
DR PFAM; PF00940; RNA_POL; 1.
KW TRANSCRIPTION; DNA-DIRECTED RNA POLYMERASE; MITOCHONDRION; PLASMID.
FT ACT_SITE 600 600 BY SIMILARITY.
FT ACT_SITE 680 680 BY SIMILARITY.
FT ACT_SITE 853 853 BY SIMILARITY.
SQ SEQUENCE 948 AA; 110728 MW; 98C32009 CRC32;

Query Match 70.08; Score 35; DB 1; Length 948;
Best Local Similarity 70.08; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKKKKKFLFL 10
Db 561 KEANKKFLFL 570
1: | | | | |
561: | | | | |

RESULT 13
RR2_CYACA
ID RR2_CYACA STANDARD; PRT; 233 AA.
AC P35014;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S2.
GN RS2.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OG CHLOROPLAST.
OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDIALES; PORPHYRIDIAEAE;
OC CYANIDIUM.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-14-1-1 / ISOLATE 107.79/GOETTINGEN;
RA KOSTRZEWA M., ZETSCHKE K.;
RT "Organization of plastid-encoded ATPase genes and flanking regions including homologues of infB and tsf in the thermophilic red alga Galdieria sulphuraria";
RL PLANT MOL. BIOL. 23:67-76(1993).
CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; X67814; G429171; -.
DR PIR; S39513; S39513.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
DR PFAM; PF00318; S2; 1.
KW RIBOSOMAL PROTEIN; CHLOROPLAST.
SQ SEQUENCE 233 AA; 26609 MW; 6C4EA9FC CRC32;

Query Match 70.08; Score 35; DB 1; Length 233;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KKKKKKFLFL 10
Db 64 KKGKKKFLFL 72
1: | | | | |
64: | | | | |

RESULT 14
Y482_METJA
ID Y482_METJA STANDARD; PRT; 240 AA.
AC Q57905;

```

```

DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0482.
GN MJ0482.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GECHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA CLOTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii";
RL SCIENCE 273:1058-1073(1996).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; U67498; G1591185; -.
DR TIGR; MJ0482; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 240 AA; 27868 MW; 4B258FF1 CRC32;

Query Match 70.08; Score 35; DB 1; Length 240;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKFL 8
Db 215 KKKKKKFL 222
1: | | | | |
215: | | | | |

RESULT 15
YCF2_MARPO
ID YCF2_MARPO STANDARD; PRT; 2136 AA.
AC P09975;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 259 KD PROTEIN (ORF 2136).
GN YCF2.
OS MARCHANTIA POLYMORPHA (LIVERWORT).
OG CHLOROPLAST.
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTES; EMBRYOPHYTES; MARCHANTIOPSIDA;
OC MARCHANTIALES; MARCHANTIAEAE; MARCHANTIA.
[1]
RN SEQUENCE FROM N.A.
RP OHYAMA K.;
RL SUBMITTED (OCT-1986) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN COMPLETE GENOME.
RA OHYAMA K., FUKUZAWA H., KOHCHI T., SHIRAI H., SANO T., SANO S.,
RA UMESONO K., SHIKI Y., TAKEUCHI M., CHANG Z., AOTA S., INOKUCHI H.,
RA OZEKI H.;
RT "Chloroplast gene organization deduced from complete sequence of liverwort Marchantia polymorpha chloroplast DNA.";
RL NATURE 322:572-574(1986).
CC -1- SIMILARITY: TO A SIMILAR ORF IN OTHER PLANTS CHLOROPLAST.

```

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X04465; G11665; .  
DR PIR; A05037; A05037.  
DR PIR; S01591; S01591.  
DR PFAM; PF00004; AAA; 1.  
DR MENDEL; 5295; MARCO.YCF2:1.  
KW CHLOROPLAST; HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 2136 AA; 239911 MW; A2FB8B07 CRC32;

Query Match 70.0%; Score 35; DB 1; Length 2136;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKKKKL 8  
| | | | |  
Db 70 KKKKKL 76

Search completed: September 8, 1999, 00:43:00  
Job time: 2615 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:54:50 ; Search time 31.14 Seconds  
(without alignments)  
12.866 Million cell updates/sec

Title: US-09-124-280A-42

Perfect score: 50  
Sequence: 1 KKKKKKFLFL 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	84.0	730	2 S54625	probable membrane
2	41	82.0	1817	2 D71606	hypothetical prote
3	40	80.0	2206	2 G71611	hypothetical prote
4	40	80.0	1182	2 G71607	probable integral
5	40	80.0	1516	2 E71619	RAD2 endonuclease
6	38	76.0	3973	2 B71612	hypothetical prote
7	37	74.0	281	2 A41400	myristylated alani
8	37	74.0	332	2 A38873	myristylated alani
9	37	74.0	335	2 S08341	myristylated alani
10	37	74.0	309	2 A39169	myristylated alani
11	37	74.0	309	2 A39773	myristylated alani
12	37	74.0	1116	2 B70476	hypothetical prote
13	36	72.0	2875	1 RRVUTW	genome polypeptin
14	36	72.0	830	2 S54547	PAM1 protein - yea
15	36	72.0	1121	2 F71613	hypothetical prote
16	36	72.0	374	2 A42264	membrane-associate
17	36	72.0	2013	2 C71610	probable membrane
18	36	72.0	154	2 S36993	transposase (clone
19	36	72.0	853	2 H70470	hypothetical prote
20	35	70.0	1386	1 RNLV2	DNA-directed RNA p
21	35	70.0	948	2 S26945	DNA-directed RNA p
22	35	70.0	233	2 S39513	ribosomal protein
23	35	70.0	100	2 S78147	ribosomal protein
24	35	70.0	85	2 T03055	zinc finger protei
25	35	70.0	270	2 E6996	amino acid ABC tra
26	35	70.0	240	2 A64360	hypothetical prote
27	35	70.0	2136	2 A05037	hypothetical prote
28	35	70.0	2539	2 B71619	hypothetical prote
29	35	70.0	1014	2 H71602	protein with DnaJ
30	35	70.0	195	2 S31861	myristylated alani
31	35	70.0	199	2 A43341	myristylated alani
32	35	70.0	200	2 S17185	myristylated alani
33	35	70.0	200	2 A47378	myristylated alani
34	34	68.0	98	1 F2NTK	photosystem II pro
35	34	68.0	326	2 F64431	capsular polysacch
36	34	68.0	417	2 F70132	conserved hypothet
37	34	68.0	435	2 H64693	fibronection/fibrin
38	34	68.0	435	2 C71807	hypothetical prote
39	34	68.0	1046	2 F71432	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S54625

probable membrane protein YOR019w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O2631; hypothetical protein YOL303.8  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 12-Dec-1997  
C:Accession: S54625; S66885  
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54617  
A:Accession: S54625  
A:Molecule type: DNA  
A:Residues: 1-730 <DEH>  
A:Cross-references: EMBL:X87331; NID:g1041652; PID:g829129  
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66877  
A:Accession: S66885  
A:Molecule type: DNA  
A:Residues: 1-730 <DEW>  
A:Cross-references: EMBL:Z74927; NID:g1420122; PID:e252323; PID:g1420123; MIPS:YOR019  
A:Experimental source: strain S288C  
C:Genetics:

A:Map position: 15R  
C:Keywords: transmembrane protein  
F:317-333/Domain: transmembrane #status predicted <TM>

Query Match 84.0% Score 42; DB 2; Length 730;

Best Local Similarity 90.0%; Pred.No. 5; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 1;

Qy 1 KKKKKKFLFL 10

||||| |||

Db 717 KKKKKKSLFL 726

##### RESULT 2

D71606

hypothetical protein PFB0800c - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: D71606

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;  
Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.  
Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600

A:Accession: D71606

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1817 <GAR>

A:Cross-references: GB:AE001418; GB:AE001362; NID:g3845275; PID:g3845280; TIGR:PFB080  
A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0800c

Query Match 82.0%; Score 41; DB 2; Length 1817;

Best Local Similarity 70.0%; Pred.No. 17;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 KKKKKKFLFL 10
    ||||| :
Db 666 KKKKKKFFYI 675

RESULT 3
G71611
hypothetical protein PFB0560w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: G71611
R: Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
  ; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: G71611
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2206 <GAR>
A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PID:g3845217; TIGR:PFB0560w
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0560w

Query Match 80.0%; Score 40; DB 2; Length 2206;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10
    ||||| :
Db 2189 KKKKKKYIYI 2198

RESULT 4
G71607
probable integral membrane protein PFB0735c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: G71607
R: Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
  ; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: G71607
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1182 <GAR>
A:Cross-references: GB:AE001414; GB:AE001362; NID:g3845260; PID:g3845263; TIGR:PFB0735c
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0735c

Query Match 80.0%; Score 40; DB 2; Length 1182;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10
    ||||| :
Db 657 KKKKKKYIYI 666

RESULT 5
E71619
RAD2 endonuclease PFB0265c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: E71619
R: Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
  ; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: E71619
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1516 <GAR>
A:Cross-references: GB:AE001383; GB:AE001362; NID:g3845135; PID:g3845138; TIGR:PFB0265c
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0265c

Query Match 80.0%; Score 40; DB 2; Length 1516;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFL 8
    |||||
Db 700 KKKKKKFL 707

RESULT 6
B71612
hypothetical protein PFB0555c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Nov-1998
C:Accession: B71612
R: Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
  ; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600
A:Accession: B71612
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-3973 <GAR>
A:Cross-references: GB:AE001402; GB:AE001362; NID:g3845213; PID:g3845215; TIGR:PFB0555c
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0555c

Query Match 76.0%; Score 38; DB 2; Length 3973;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10
    ||||| :
Db 2630 KKKKKKYIFL 2639

RESULT 7
A41400
myristylated alanine-rich protein kinase C substrate - chicken
C:Species: Gallus gallus (chicken)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 13-Mar-1998
C:Accession: A41400
R: Graff, J.M.; Stumpo, D.J.; Blackshear, P.J.
Mol. Endocrinol. 3, 1903-1906, 1989
A:Title: Molecular cloning, sequence, and expression of a cDNA encoding the chicken m
A:Reference number: A41400; MUID:90114197
A:Accession: A41400
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-281 <GRA>
A:Cross-references: GB:M31650; NID:g212287; PID:g212288
C:Keywords: actin binding; phosphoprotein

Query Match 74.0%; Score 37; DB 2; Length 281;
Best Local Similarity 77.8%; Pred. NO. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```



Qy 1 KKKKKKFLF 9  
 |||||:|  
 Db 117 KKKKKRFSF 125

RESULT 8  
 A38873  
 myristylated alanine-rich protein kinase C substrate - human  
 N;Alternate names: acidic calmodulin-binding 80K protein; MARCKS  
 C;Species: Homo sapiens (man)  
 C;Date: 30-Sep-1993 #sequence.revision 30-Sep-1993 #text\_change 05-Feb-1999  
 C;Accession: A38873; A42977; A40758; S29269  
 R;Shimizu, N.  
 submitted to DBJ, September 1991  
 A;Reference number: A38873  
 A;Accession: A38873  
 A;Molecule type: mRNA  
 A;Residues: 1-332 <SHI>  
 A;Cross-references: GB:D10522; GB:D90498; NID:g219893; PID:d1001865; PID:g219894  
 R;Sakai, K.; Hirai, M.; Kudoh, J.; Minoshima, S.; Shimizu, N.  
 Genomics 14, 175-178, 1992  
 A;Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L protein  
 A;Reference number: A42977; MUID:93052291  
 A;Accession: A42977  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-13, 'G', 15-332 <SAK>  
 A;Cross-references: GB:D90498  
 A;Experimental source: squamous carcinoma cells A431  
 A;Note: sequence extracted from NCBI backbone (NCBIP:118653)  
 R;Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.; Bla  
 J. Biol. Chem. 266, 14399-14405, 1991  
 A;Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACS). A  
 A;Reference number: A40758; MUID:91317795  
 A;Accession: A40758  
 A;Molecule type: mRNA  
 A;Residues: 1-83, 'A', 85-118, 'P', 120-233, 'W', 235-286, 'LVC', 290, 'RRGGSPGGARGRRSLNQ', 309-3  
 A;Cross-references: GB:M68956  
 A;Note: the authors translated the codon GGC for residue 53 as Arg  
 R;Herget, T.; Brooks, S.F.; Broad, S.; Rozenfurt, E.  
 Eur. J. Biochem. 209, 7-14, 1992  
 A;Title: Relationship between the major protein kinase C substrates acidic 80-kDa protei  
 or equivalent genes in different species.  
 A;Reference number: S29267; MUID:93011168  
 A;Accession: S29269  
 A;Molecule type: protein  
 A;Residues: 2-11 <MAN>  
 R;Mizutani, A.; Tokumitsu, H.; Hidaka, H.  
 Biochem. Biophys. Res. Commun. 182, 1395-1401, 1992  
 A;Title: Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein interacting w  
 A;Reference number: PS0338; MUID:92171958  
 A;Accession: PS0338  
 A;Molecule type: protein  
 A;Residues: 12-30;56-69;88-98, 'AS', 100-103;104-109, 'E', 111-123;156-160;165-171;196-21  
 A;Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myrist  
 F;2/Modified site: myristoylated amino end (Gly) (in mature form) #status predicted  
 C;Comment: This protein is a major cellular substrate for protein kinase C and plays  
 C;Comment: It binds to calmodulin in one to one molar ratio in the presence of calciu  
 C;Genetics:  
 A;Introns: 34/3  
 Query Match 74.0%; Score 37; DB 2; Length 335;  
 Best Local Similarity 77.8%; Pred. No. 17;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KKKKKKFLF 9  
 |||||:|  
 Db 151 KKKKKRFSF 159

RESULT 10  
 A39169  
 myristylated alanine-rich protein kinase C substrate, macrophage - mouse  
 N;Alternate names: MARCKS; myristoylated alanine-rich C kinase substrate  
 C;Species: Mus musculus (house mouse)  
 C;Date: 24-Jan-1992 #sequence.revision 24-Jan-1992 #text\_change 05-Dec-1998  
 C;Accession: A39169; S16519; S29268; S63977; S74153  
 R;Seykora, J.T.; Ravetch, J.V.; Aderem, A.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 2505-2509, 1991  
 A;Title: Cloning and molecular characterization of the murine macrophage "68-kDa" pro  
 A;Reference number: A39169; MUID:91172836

Qy 1 KKKKKKFLF 9  
 |||||:|  
 Db 152 KKKKKRFSF 160

RESULT 9  
 S08341  
 myristylated alanine-rich protein kinase C substrate - bovine  
 N;Alternate names: acidic calmodulin-binding 81K protein (ACAMP-81); MARCKS  
 C;Species: Bos primigenius taurus (cattle)

A:Accession: A39169  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-309 <SEY>  
 A:Cross-references: GB:M60474; NID:g199026; PID:g199027  
 R:Brooks, S.F.; Herget, T.; Erusalimsky, J.D.; Rozengurt, E.  
 EMBO J. 10, 2497-2505, 1991  
 A:Title: Protein kinase C activation potentially down-regulates the expression of its major  
 A:Reference number: S16519; MUID:91330872  
 A:Accession: S16519  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-95, 'T', '97', 'T', '99-309 <EMB>  
 R:Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
 Eur. J. Biochem. 209, 7-14, 1992  
 A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein  
 or equivalent genes in different species.  
 A:Reference number: S29267; MUID:93011168  
 A:Accession: S29268  
 A:Molecule type: mRNA  
 A:Residues: 182-301 <HER>  
 R:Herget, T.; Oehlelein, S.A.; Pappin, D.J.C.; Rozengurt, E.; Parker, P.J.  
 Eur. J. Biochem. 233, 448-457, 1995  
 A:Title: The myristoylated alanine-rich C-kinase substrate (MARCKS) is sequentially phos  
 A:Reference number: S63977  
 A:Accession: S63977  
 A:Molecule type: protein  
 A:Residues: 145-151, 'X', '153-155, 'X', '157-162, 'X', '164-169 <HEW>  
 R:Schoenwasser, D.C.; Palmer, R.H.; Herget, T.; Parker, P.J.  
 FEBS Lett. 395, 1-5, 1996  
 A:Title: p42 MAPK phosphorylates 80 kDa MARCKS at Ser-113.  
 A:Reference number: S74153  
 A:Accession: S74153  
 A:Molecule type: protein  
 A:Residues: 102-130 <SCH>  
 A:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyla  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:152,156,160,163/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status  
 Query Match 74.0%; Score 37; DB 2; Length 309;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKKKKKFLF 9  
 |||||:|  
 Db 145 KKKKKRFSF 153  
 RESULT 11  
 A39773  
 myristoylated alanine-rich protein kinase C substrate, macrophage - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 13-Mar-1998  
 C:Accession: A39773  
 R:Erusalimsky, J.D.; Brooks, S.F.; Herget, T.; Morris, C.; Rozengurt, E.  
 J. Biol. Chem. 266, 7073-7080, 1991  
 A:Title: Molecular cloning and characterization of the acidic 80-kDa protein kinase C su  
 A:Reference number: A39773; MUID:91201362  
 A:Accession: A39773  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-309 <ERU>  
 A:Cross-references: GB:M59859  
 C:Keywords: phosphoprotein  
 Query Match 74.0%; Score 37; DB 2; Length 309;  
 Best Local Similarity 77.8%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKKKKKFLF 9  
 |||||:|

Db 145 KKKKKRFSF 153  
 RESULT 12  
 B70476  
 hypothetical protein aq\_2054 - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998  
 C:Accession: B70476  
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
 V.  
 Nature 392, 353-358, 1998  
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666  
 A:Accession: B70476  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1116 <AQF>  
 A:Cross-references: GB:AE000770; NID:g2984274; PID:g2984280; GB:AE000657  
 A:Experimental source: strain VF5  
 C:Genetics:  
 A:Gene: aq\_2054  
 Query Match 74.0%; Score 37; DB 2; Length 1116;  
 Best Local Similarity 87.5%; Pred. No. 53;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KKKKKKFL 8  
 :|||||  
 Db 84 KKKKKKFL 91  
 RESULT 13  
 RRVUTW  
 genome polyprotein - tomato spotted wilt virus (strain BR-01)  
 N:Alternate names: L protein  
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
 C:Species: tomato spotted wilt virus  
 C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
 C:Accession: JQ1335  
 R:de Haan, P.; Kormelink, R.; de Oliveira Resende, R.; van Poelwijk, F.; Peters, D.;  
 J. Gen. Virol. 71, 2207-2216, 1991  
 A:Title: Tomato spotted wilt virus L RNA encodes a putative RNA polymerase.  
 A:Reference number: JQ1335  
 A:Accession: JQ1335  
 A:Molecule type: genomic RNA  
 A:Residues: 1-2875 <DEH>  
 C:Genetics:  
 A:Map position: segment L  
 C:Superfamily: bunyavirus RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase  
 Query Match 72.0%; Score 36; DB 1; Length 2875;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKKKKKFLF 9  
 :|||||  
 Db 2669 KKKKKGFLF 2677  
 RESULT 14  
 S54547  
 FAM1 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YD8419.18; protein YD9320A.01; protein YDR251w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 23-Aug-1996 #text\_change 06-Feb-1998  
 C:Accession: S54547; A53049; S67308; S67454; S45179  
 R:Oliver, K.; Harris, D.  
 Submitted to the EMBL Data Library, May 1995  
 A:Reference number: S54530

Search completed: September 7, 1999, 23:54:51  
Job time: 1902 sec

A:Accession: S54547  
A:Molecule type: DNA  
A:Residues: 1-659 <OLI>  
A:Cross-references: EMBL:Z49701; NID:g817819; PID:g817837; MIPS:YDR251w  
R:Hu, G.Z.; Ronne, H.  
J. Biol. Chem. 269, 3429-3435, 1994  
A:Title: Overexpression of yeast PAM1 gene permits survival without protein phosphatase  
A:Reference number: A53049; MUID:94148839  
A:Accession: A53049  
A:Molecule type: DNA  
A:Residues: 1-255, 'T', 257-830 <HUA>  
A:Cross-references: EMBL:X73454; NID:g456666; PID:g456667  
R:Murphy, L.; Harris, D.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: S61117  
A:Accession: S67308  
A:Molecule type: DNA  
A:Residues: 660-830 <MUR>  
A:Cross-references: EMBL:Z68329; NID:g1136205; PID:e215297; PID:g1136206; MIPS:YDR251w  
R:Murphy, L.; Harris, D.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: S67454  
A:Accession: S67454  
A:Molecule type: DNA  
A:Residues: 660-830 <MUW>  
A:Cross-references: EMBL:Z70202; NID:g1226026; PID:e228706; PID:g1226027  
C:Genetics:  
A:Gene: SGD:PAM1  
A:Cross-references: SGD:S0002659; MIPS:YDR251w  
A:Map position: 4R  
C:Keywords: coiled coil

Query Match 72.0%; Score 36; DB 2; Length 830;  
Best Local Similarity 77.8%; Pred. No. 60;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
:|||||  
Db 816 EKKKKKFSF 824

RESULT 15  
F71613  
hypothetical protein PFB0495w - malaria parasite (Plasmodium falciparum)  
C:Species: plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 13-Nov-1998  
C:Accession: F71613  
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600  
A:Accession: F71613  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1121 <GAR>  
A:Cross-references: GB:AE001398; GB:AE001362; NID:g3845197; PID:g3845199; TIGR:PFB0495w  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0495w

Query Match 72.0%; Score 36; DB 2; Length 1121;  
Best Local Similarity 77.8%; Pred. No. 79;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
:|||||  
Db 1047 KMEKKKFLF 1055

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:23:07 ; Search time 25.2 Seconds  
(without alignments)  
3.916 Million cell updates/sec

Title: US-09-124-280A-42  
Perfect score: 50  
Sequence: 1 KKKKKKFLFL 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	10	2	US-08-456-112B-42
2	40	80.0	10	1	US-08-097-830E-12
3	40	80.0	10	2	US-08-456-112B-12
4	37	74.0	332	2	US-08-405-175A-5
5	37	74.0	335	2	US-08-405-175A-6
6	37	74.0	309	2	US-08-405-175A-7
7	37	74.0	309	2	US-08-405-175A-8
8	37	74.0	281	2	US-08-405-175A-9
9	33	66.0	539	2	US-08-978-182-3
10	32	64.0	10	1	US-08-097-830E-19
11	32	64.0	319	1	US-08-597-236-7
12	32	64.0	319	2	US-08-746-682A-7
13	32	64.0	627	2	US-08-703-947-2
14	32	64.0	10	2	US-08-456-112B-19
15	32	64.0	63	3	PCT-US94-01202-2
16	31	62.0	473	1	US-08-103-739B-2
17	31	62.0	772	1	US-08-258-639A-2
18	31	62.0	182	2	US-08-226-264-28
19	31	62.0	473	2	US-08-47A-404-2
20	31	62.0	594	2	US-08-659-939-2
21	31	62.0	527	2	US-08-659-939-4
22	31	62.0	473	2	US-08-485-845-2
23	31	62.0	473	2	US-08-482-714-2
24	31	62.0	594	2	US-08-850-041-2
25	31	62.0	527	2	US-08-850-041-4
26	31	62.0	572	3	PCT-US91-08177-11
27	31	62.0	558	3	PCT-US93-03027-2
28	31	62.0	772	3	PCT-US95-07391A-2
29	31	62.0	10	4	521624-22
30	30	60.0	888	1	US-07-718-575-6
31	30	60.0	902	1	US-07-718-575-8
32	30	60.0	269	1	US-08-203-716-3
33	30	60.0	11	1	US-07-694-983-15
34	30	60.0	561	1	US-08-096-947-1
35	30	60.0	501	1	US-08-190-802A-28
36	30	60.0	576	1	US-08-190-802A-56
37	30	60.0	510	1	US-08-249-112-3
38	30	60.0	487	1	US-08-249-112-4
39	30	60.0	1213	1	US-08-188-582-20

40 30 60.0 32 1 US-08-152-488-8 Sequence 8, Appl1  
41 30 60.0 29 1 US-08-054-363-17 Sequence 17, Appl1  
42 30 60.0 36 1 US-07-948-357-4 Sequence 4, Appl1  
43 30 60.0 36 1 US-07-948-357-12 Sequence 12, Appl1  
44 30 60.0 36 1 US-07-948-357-13 Sequence 13, Appl1  
45 30 60.0 37 1 US-08-231-730A-29 Sequence 23, Appl1

ALIGNMENTS

RESULT 1  
US-08-456-112B-42  
; Sequence 42, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-42  
  
Query Match 100.0%; Score 50; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.047;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10  
Db 1 KKKKKKFLFL 10

RESULT 2  
US-08-097-830E-12  
; Sequence 12, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porro, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097.830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-12

Query Match 80.0%; Score 40; DB 1; Length 10;  
Best Local Similarity 88.9%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
Db 1 KKKKKKFKF 9

RESULT 3  
US-08-456-112B-12  
Sequence 12, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: POTO, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-12

Query Match 80.0%; Score 40; DB 2; Length 10;  
Best Local Similarity 88.9%; Pred. No. 1.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
Db 1 KKKKKKFKF 9

RESULT 4  
US-08-405-175A-5  
Sequence 5, Application US/08405175A  
Patent No. 5885772  
GENERAL INFORMATION:  
APPLICANT: Aderem, Alan A.  
APPLICANT: Chen, Jianmin  
APPLICANT: Chang, Sandy  
TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,175A  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-121A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 332 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: predicted primary structure of human MARCKS  
HYPOTHETICAL: NO  
US-08-405-175A-5

Query Match 74.0%; Score 37; DB 2; Length 332;  
Best Local Similarity 77.8%; Pred. No. 46;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
Db 152 KKKKKRFSF 160

```
RESULT 5
US-08-405-175A-6
; Sequence 6, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405.175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: predicted primary structure of bovine MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-6

Query Match 74.0%; Score 37; DB 2; Length 335;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKELF 9
Db 151 KKKKKRFSF 159

RESULT 6
US-08-405-175A-7
; Sequence 7, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405.175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: predicted primary structure of bovine MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-7

Query Match 74.0%; Score 37; DB 2; Length 335;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKELF 9
Db 151 KKKKKRFSF 159
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405.175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: predicted primary structure of murine MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-7

Query Match 74.0%; Score 37; DB 2; Length 309;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKELF 9
Db 145 KKKKKRFSF 153

RESULT 7
US-08-405-175A-8
; Sequence 8, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405.175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
```

SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: predicted primary structure of rat MARCKS  
HYPOTHETICAL: NO  
US-08-405-175A-8

Query Match 74.0%; Score 37; DB 2; Length 309;  
Best Local Similarity 77.8%; Pred. NO. 41;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
DB 145 KKKKKRFSF 153

RESULT 8  
US-08-405-175A-9  
; Sequence 9, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianlin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405.175A  
; FILING DATE: 16-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-121A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 281 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: predicted primary structure of chicken MARCKS  
; HYPOTHETICAL: NO  
US-08-405-175A-9

Query Match 74.0%; Score 37; DB 2; Length 281;  
Best Local Similarity 77.8%; Pred. NO. 41;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
DB 117 KKKKKRFSF 125

RESULT 9  
US-08-978-182-3  
; Sequence 3, Application US/08978182  
; Patent No. 5849556  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Kaser, Matthew  
; APPLICANT: Mathur, Preete  
; TITLE OF INVENTION: HUMAN GROWTH-RELATED CDC10 HOMOLOG  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,182  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0426 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 539 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 508229  
US-08-978-182-3

Query Match 66.0%; Score 33; DB 2; Length 539;  
Best Local Similarity 77.8%; Pred. NO. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
DB 531 KKKKKGLF 539

RESULT 10  
US-08-097-830E-19  
; Sequence 19, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: Porzio, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.



STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:  
OTHER INFORMATION: sulfide bond between Cys and Cys  
US-08-097-830E-19

Query Match 64.0%; Score 32; DB 1; Length 10;  
Best Local Similarity 85.7%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KKKKFLF 9  
|||||  
Db 2 KKKKFF 8

RESULT 11  
US-08-597-236-7  
Sequence 7, Application US/08597236  
Patent No. 5733765  
GENERAL INFORMATION:  
APPLICANT: STINGELE, Francesca  
APPLICANT: MOLLET, Beat  
TITLE OF INVENTION: LACTIC BACTERIA PRODUCING  
TITLE OF INVENTION: EXOPOLYSACCHARIDES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/597,236  
FILING DATE:  
CLASSIFICATION: 426  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95201669.9  
FILING DATE: 20-JUN-1995  
ATTORNEY/AGENT INFORMATION:

NAME: Fanucci A., Allan  
REGISTRATION NUMBER: 30256  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 319 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-597-236-7

Query Match 64.0%; Score 32; DB 1; Length 319;  
Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 KKKFLFL 10  
|||||  
Db 81 KKKFLFV 87

RESULT 12  
US-08-746-682A-7  
Sequence 7, Application US/08746682A  
Patent No. 5786184  
GENERAL INFORMATION:  
APPLICANT: STINGELE, Francesca  
APPLICANT: MOLLET, Beat  
TITLE OF INVENTION: LACTIC BACTERIA PRODUCING  
TITLE OF INVENTION: EXOPOLYSACCHARIDES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/746,682A  
FILING DATE: 14-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/597,236  
FILING DATE: 20-JUN-1995  
APPLICATION NUMBER: EP 95201669.9  
FILING DATE: 20-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fanucci A., Allan  
REGISTRATION NUMBER: 30256  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 319 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-746-682A-7

Query Match 64.0%; Score 32; DB 2; Length 319;  
Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 KKKFLFL 10  
|||||  
Db 81 KKKFLFV 87

RESULT 13

```

; LOCATION: entire derived coded sequence
; IDENTIFICATION METHOD: clone identified by immunodetection of
; IDENTIFICATION METHOD: protein product with antiserum specific for
; IDENTIFICATION METHOD: P65; residue sequence deduced from nucleic
; IDENTIFICATION METHOD: acid sequence
; OTHER INFORMATION: immunogenic surface lipoprotein of no known
; OTHER INFORMATION: function: C-terminus exposed on external
; OTHER INFORMATION: surface of cell; N-terminal signal sequence
; OTHER INFORMATION: (first 29 amino acid residues) cleaved during
; OTHER INFORMATION: lipid modification process
; PUBLICATION INFORMATION:
; AUTHORS: Mary F. Kim, Manijeh B. Heidari, Susan J.
; TITLES: Identification and Mapping of an
; IMMUNOGENIC REGION OF MYCOPLASMA
; HYOPNEUMONIAE P65 SURFACE LIPOPROTEIN
; TITLE: Expressed in Escherichia coli from a Cloned
; TITLE: Genomic Fragment
; JOURNAL: Infection and Immunity
; VOLUME: 58
; ISSUE: 8
; PAGES: 2637-2643
; DATE: August 1990
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: From 1 to 627
; US-08-703-947-2

Query Match 64.0%; Score 32; DB 2; Length 627;
Best Local Similarity 70.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps

QY 1 KKKKKKFLFL 10
   ||| :||| |
DB 2 KKKARKFLRL 11

RESULT 14
US-08-456-112B-19
; Sequence 19 Application US/08456112B
; Patent No. 5834430
; GENERAL INFORMATION:
; APPLICANT: POTO, Massimo
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hedman, Gibson & Costigan
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: LEADING EDGE 486
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,112B
; FILING DATE: May 31, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Costigan, James V.
; REGISTRATION NUMBER: 25,669
; REFERENCE/DOCKET NUMBER: 576-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 302-8989
; TELEFAX: (212) 302-8998

```

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular

US-08-456-112B-19

Query Match 64.0%; Score 32; DB 2; Length 10;  
Best Local Similarity 85.7%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KKKKFLF 9  
Db 2 KKKKFFF 8

RESULT 15

PCT-US94-01202-2  
Sequence 2, Application PC/TUS9401202

GENERAL INFORMATION:

APPLICANT: Ribeiro, Jose M.  
APPLICANT: Lerner, Ethan A.  
APPLICANT: Remold, Heinz G.  
APPLICANT: Titus, Richard G.  
APPLICANT: Tsuji, Yoshiharu  
APPLICANT: Hansawa, Chika  
APPLICANT: Uzuka, Makoto

TITLE OF INVENTION: Method for Enhancing Hair Growth and  
TITLE OF INVENTION: Hair Revitalizing Compositions Comprising Lutzomyia

TITLE OF INVENTION: Protein

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/01202

FILING DATE: 10-FEB-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/017,061

FILING DATE: 12-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 05136-0002-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 63 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US94-01202-2

Query Match 64.0%; Score 32; DB 3; Length 63;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKKKF 7

Db 53 KKKKKKF 59

Search completed: September 7, 1999, 23:23:07  
Job time: 911 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:18:19 ; Search time 40.95 Seconds  
(without alignments)  
5.784 Million cell updates/sec

Title: US-09-124-280A-42  
Perfect score: 50  
Sequence: 1 KKKKKKFLFL 10  
Scoring table: BLOSUM62  
Searched: 188963 seqs, 23686106 residues  
Database : A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	100.0	10	1	W21630	Antibiotic potenti
2	41	82.0	29	1	W22829	MICFTR based chann
3	41	82.0	30	1	W22830	MICFTR based chann
4	40	80.0	34	1	R11514	Anti-ATLA antibody
5	40	80.0	10	1	R71783	Peptide neutralisi
6	40	80.0	10	1	W21600	Antibiotic potenti
7	37	74.0	330	1	R05528	High density lipop
8	37	74.0	20	1	W45880	Peptide membrane b
9	37	74.0	25	1	W71666	MARCKS-derived tar
10	37	74.0	30	1	W86744	Anticoagulant pept
11	36	72.0	13	1	W24440	Nucleic acid (NA)
12	36	72.0	28	1	W22828	MICFTR based chann
13	35	70.0	21	1	R05502	Modified signal se
14	35	70.0	584	1	W12661	Orange pectin meth
15	35	70.0	24	1	W71667	MacMARCKS-derived
16	35	70.0	584	1	W72964	Aspergillus aculea
17	34	68.0	22	1	R05503	Modified signal se
18	34	68.0	317	1	R21750	C-terminal sequenc
19	34	68.0	474	1	R32587	P.falciiparum antig
20	34	68.0	474	1	R32992	P.falciiparum antig
21	34	68.0	13	1	R57406	Peptide for treati
22	34	68.0	417	1	R70649	Mouse azoospermia
23	34	68.0	363	1	W20581	Helicobacter pylor
24	34	68.0	363	1	W24696	H. pylori cell env
25	34	68.0	200	1	W44087	Human secreted pro
26	34	68.0	200	1	W27651	Secreted protein A
27	34	68.0	436	1	W55473	H. pylori ORF 06cp
28	34	68.0	436	1	W55543	H. pylori ORF 06cp
29	34	68.0	442	1	W55250	H. pylori ORF 05cp
30	33	66.0	29	1	R11513	Anti-ATLA antibody
31	33	66.0	152	1	R15746	Interleukin-1 beta
32	33	66.0	481	1	R33389	AII/AVPv2 receptor
33	33	66.0	35	1	W38232	K6CLII peptide use
34	33	66.0	35	1	W65932	Peptide K6CL22 use
35	33	66.0	35	1	W65933	Peptide CL28 used
36	33	66.0	41	1	W65934	Peptide CL26 used
37	33	66.0	40	1	W65939	Polylysine peptide
38	33	66.0	22	1	W65940	Polylysine peptide
39	33	66.0	49	1	W59573	Immunoglobulin IgG
40	32	64.0	143	1	R14214	HCV detecting pept
41	32	64.0	38	1	R14210	HCV detecting pept
42	32	64.0	179	1	R14216	HCV detecting pept
43	32	64.0	143	1	R14217	HCV detecting pept

ALIGNMENTS

RESULT 1

W21630 ID W21630 standard; peptide; 10 AA.  
AC W21630:  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #42.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
PN W09638163-A1.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 46; Page 29; 37pp; English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 50; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLFL 10  
| | | | | | | | | |  
Db 1 KKKKKKFLFL 10

RESULT 2

W22829 ID W22829 standard; peptide; 29 AA.  
AC W22829:  
DT 13-MAR-1998 (first entry)  
DE MICFTR based channel-forming peptide 27.  
KW Channel-forming peptide; channel assembly; epithelial cell; treatment;  
KW cystic fibrosis; polycystic kidney disease; anion transportation; MICFTR.  
OS Synthetic.  
PN Homo sapiens.  
PD W09726905-A1.  
PF 31-JUL-1997. U01103.  
PR 27-JAN-1997; US-789155.  
PR 24-JAN-1997; US-789155.  
PR 25-JAN-1996; US-591381.  
PR 23-JAN-1997; US-591381.  
PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
PA (UNIV ) UNIV KANSAS STATE RES FOUND.  
PI Iwamoto T, Sullivan LP, Tomich JM;  
DR WPI; 97-393366/36.  
PT Channel assembly for transporting ions across epithelial cell  
PT membranes - comprises new water soluble peptide(s), for treating  
PT cystic fibrosis and polycystic kidney disease by altering water flux  
PT across cells  
PS Example 13; Page 64; 93pp; English.  
CC This MICFTR based peptide is comprised of hydrophobic residues typical of

44 32 64.0 38 1 R14775 HCV detecting pept  
45 32 64.0 319 1 W14075 S.thermophilus exo

CC a transmembrane sequence. The M1CFTR peptide associates with membranes  
 CC but does not form an ionic conducting channel. This peptide, however with  
 CC additional lysine residues at the N-terminal, can be used to construct a  
 CC novel channel assembly comprising 3-6 novel peptides, of 18-30 amino  
 CC acids. Suitable protein fragments for use as such peptides are present in  
 CC the strychnine-binding alpha-subunit of the inhibitory glycine receptor  
 CC (M2GlyR) from human brain, and the inhibitory gamma-aminobutyric acid  
 CC receptor from human brain, and the cystic fibrosis transmembrane  
 CC conductance regulator from human epithelium. The peptides are synthesized  
 CC by standard solid phase peptide synthesis. The M2GlyR based peptides are  
 CC water soluble to at least 10 mM and enables anions to be transported  
 CC through a membrane of an epithelial cell when they are embedded in the  
 CC membrane. The channel assembly can be used to alter the flux of water  
 CC across an epithelial cell, particularly for treatment of cystic fibrosis  
 CC (where affected cells are in the airway, pancreatic duct or epididymis).  
 CC The channel assembly can also be used in the treatment of autosomal  
 CC dominant polycystic kidney disease (where the affected cells are in the  
 CC cystic epithelium). The channel assembly spontaneously inserts into the  
 CC basolateral membrane to prevent water flow to adjacent cysts.  
 CC Sequence 29 AA;

Query Match 82.0%; Score 41; DB 1; Length 29;  
 Best Local Similarity 77.8%; Pred. No. 1.1;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
 Db 1 KKKKKRFFM 9

RESULT 3  
 W2830  
 ID W2830 standard; peptide; 30 AA.

DT 13-MAR-1998 (first entry)  
 DE M1CFTR based channel-forming peptide 28.  
 KW Channel-forming peptide; channel assembly; epithelial cell; treatment;  
 OS cystic fibrosis; polycystic kidney disease; anion transportation; M1CFTR.  
 OS Synthetic.  
 OS Homo sapiens.  
 PN W09726905-A1.  
 PD 31-JUL-1997.  
 PF 27-JAN-1997; U01103.  
 PR 24-JAN-1997; US-789155.  
 PR 25-JAN-1996; US-591381.  
 PR 23-JAN-1997; US-591381.  
 PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
 PA (UNIV ) UNIV KANSAS STATE RES FOUND.  
 PI Iwanoto T, Sullivan LP, Tomich JM;  
 DR WPI: 97-393366/36.

PT Channel assembly for transporting ions across epithelial cell  
 PT membranes - comprises new water soluble peptide(s); for treating  
 PT cystic fibrosis and polycystic kidney disease by altering water flux  
 PT across cells

PS Example 13; Page 65; 93pp; English.  
 CC This M1CFTR based peptide is comprised of hydrophobic residues typical of  
 CC a transmembrane sequence. The M1CFTR peptide associates with membranes  
 CC but does not form an ionic conducting channel. This peptide, however with  
 CC additional lysine residues at the N-terminal, can be used to construct a  
 CC novel channel assembly comprising 3-6 novel peptides, of 18-30 amino  
 CC acids. Suitable protein fragments for use as such peptides are present in  
 CC the strychnine-binding alpha-subunit of the inhibitory glycine receptor  
 CC (M2GlyR) from human brain, the inhibitory gamma-aminobutyric acid  
 CC receptor from human brain, and the cystic fibrosis transmembrane  
 CC conductance regulator from human epithelium. The peptides are synthesized  
 CC by standard solid phase peptide synthesis. The M2GlyR based peptides are  
 CC water soluble to at least 10 mM and enables anions to be transported  
 CC through a membrane of an epithelial cell when they are embedded in the  
 CC membrane. The channel assembly can be used to alter the flux of water  
 CC across an epithelial cell, particularly for treatment of cystic fibrosis  
 CC (where affected cells are in the airway, pancreatic duct or epididymis).  
 CC The channel assembly can also be used in the treatment of autosomal

CC dominant polycystic kidney disease (where the affected cells are in the  
 CC cystic epithelium). The channel assembly spontaneously inserts into the  
 CC basolateral membrane to prevent water flow to adjacent cysts.  
 CC Sequence 30 AA;

Query Match 82.0%; Score 41; DB 1; Length 30;  
 Best Local Similarity 77.8%; Pred. No. 1.1;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKKKKKFLF 9  
 Db 2 KKKKKRFFM 10

RESULT 4

R11514  
 ID R11514 standard; Protein; 34 AA.  
 AC R11514;  
 DT 21-JUN-1991 (first entry)  
 DE Anti-ATLA antibody-binding peptide #3.  
 KW adult T cell leukaemia associated antigen; HTLV-1; ATL.  
 OS Synthetic.  
 PN EP-423649-A.  
 PD 24-APR-1991.  
 PF 12-OCT-1990; 119624.  
 PR 13-OCT-1989; JP-266983.  
 PA (KURS ) KURARAY KK.  
 PI Maeda Y, Shiraki H, Washitani Y, Kuroda N, Yamada K;  
 PI Oka K, Namba T;  
 DR WPI: 91-118942/17.  
 PT New peptide(s) able to bind to anti-ATLA antibody - used in  
 PT diagnosis and to treat HTLV-1-associated myelopathy and diseases  
 PT caused by ATL.  
 PS Claim 4; Page 31; 44 pp; English.  
 CC This peptide is able to bind to an antibody specific against an  
 CC adult T cell leukaemia associated antigen. The peptide can have  
 CC from 1 to 10 Lys residues at the N-terminus. The invention also  
 CC covers an adsorbent for anti-ATLA antibodies, comprising the peptide  
 CC immobilised on a carrier and a reagent containing the peptide, for  
 CC measuring the antibody. The reagent and the adsorbent can be used  
 CC to treat an adult T-cell leukaemia virus-infectious disease.  
 CC See also R11512 and R11513.  
 CC Sequence 34 AA;

Query Match 80.0%; Score 40; DB 1; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKKKKKFL 8  
 Db 5 KKKKKKFL 12

RESULT 5

R71783  
 ID R71783 standard; peptide; 10 AA.  
 AC R71783;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.  
 PN W09503327-A.  
 PD 02-FEB-1995.  
 PF 21-JUL-1994; E02413.  
 PR 26-JUL-1993; US-097830.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M;  
 DR WPI: 95-075190/10.  
 PT New peptide(s) for neutralising LPS endotoxin - comprising  
 PT repeating units of a basic amino acid or basic and hydrophobic  
 PT amino acids

PS Claim 14; Page 21; 26pp; English.  
 CC New peptides are claimed which are linear or cyclic peptides which  
 CC include units of formula: (A)<sub>n</sub>, where A is the cationic amino acid Lys  
 CC or Arg and n is 7-10; (AB)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic  
 CC amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and  
 CC (ABC)<sub>p</sub>, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or  
 CC Trp, and p is 2 or greater.  
 CC The peptides bind to lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 80.0%; Score 40; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKFLF 9  
 |||||  
 Db 1 KKKKKKFKF 9

## RESULT 6

W21600  
 ID W21600 standard; peptide; 10 AA.  
 AC W21600;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #12.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 16; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 80.0%; Score 40; DB 1; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.59;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKFLF 9  
 |||||  
 Db 1 KKKKKKFKF 9

## RESULT 7

R05528  
 ID R05528 standard; protein; 330 AA.  
 AC R05528;  
 DT 23-OCT-1990 (first entry)  
 DE High density lipoprotein (HDL) binding protein.  
 KW High density lipoprotein; HDL-binding protein; atherosclerosis;  
 KW hypercholesterolaemia; ds.

OS Homo sapiens.  
 PN W09005744-A.  
 PD 31-MAY-1990.  
 PF 17-NOV-1989; 005169.  
 PR 18-NOV-1988; US-273388,  
 PA (UNIM) Univ of Washington, (ZYMO-) Zymogenetics Inc.  
 PI Oram JF, McKnight GL, Hart CE, Curtis DA;  
 DR WPI: 90-193405/25.  
 DR N-PSDB; Q04784.  
 DT New mammalian proteins binding high density lipoprotein sub-class 3 -  
 DT DNA encoding them and derived antibodies, for screening  
 DT potentially therapeutic HDL analogues and for diagnosing risk of  
 DT atherosclerosis.  
 PS Claim 4; Fig 1A-D; 79pp; English.  
 CC The protein product may be used to raise Abs, and the CDNA to  
 CC create probes, both useful in screening for HDL analogues,  
 CC agonists and antagonists, and in identifying abnormalities in the  
 CC HDL binding/receptor pathway. HDL analogues can be used in treating  
 CC hypercholesterolaemia and atherosclerosis  
 SQ Sequence 330 AA;

Query Match 74.0%; Score 37; DB 1; Length 330;  
 Best Local Similarity 77.8%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKKKKKFLF 9  
 |||||  
 Db 152 KKKKKKFSF 160

## RESULT 8

W45880  
 ID W45880 standard; peptide; 20 AA.  
 AC W45880;  
 DT 30-JUN-1998 (first entry)  
 DE Peptide membrane binding element.  
 KW Membrane binding element; thrombotic disease; inflammation;  
 KW complement related disease; soluble peptide.  
 OS Synthetic.  
 PN W09802454-A2.  
 PD 22-JAN-1998.  
 PF 08-JUL-1997; E03715.  
 PR 15-JUL-1996; GB-014871.  
 PA (ADPR-) ADPROTECH PLC.  
 PI Dodd I, Mossakowska DEI, Smith RAG;  
 DR WPI: 98-110524/10.  
 PT Derivatives of soluble poly(peptide(s) bonded to low affinity  
 PT membrane binding groups - useful for treating complement-related and  
 PT thrombotic diseases, providing improved localisation at cellular  
 PT membranes  
 PS Claim 11; Page 70; 75pp; English.  
 CC The present peptide sequence represents a specifically claimed membrane  
 CC binding element. The invention relates to a soluble derivative (A) of a  
 CC soluble polypeptide (I), which comprises at least 2 heterologous  
 CC membrane-binding elements (MBE) of low membrane affinity covalently  
 CC associated with (I). MBE interact, independently and with thermodynamic  
 CC additivity, with components of cellular or artificial membranes exposed  
 CC to extracellular fluids. (A) are used to treat disorders treatable with  
 CC (I) itself, specifically inflammation or any other complement-related  
 CC disorder (e.g. neurological disease, graft rejection, myocardial  
 CC infarction, sepsis, rheumatoid arthritis and many others; including  
 CC application to indwelling devices) and thrombolytic disease, but also to  
 CC treat allergy, induce weight loss, to treat ischaemia or asthma and as  
 CC immuno-modulators for treating multiple sclerosis. (A) are administered  
 CC orally, topically, by injection or inhalation at 0.01-10 (preferably  
 CC 0.1-10) mg/kg/day.  
 SQ Sequence 20 AA;

Query Match 74.0%; Score 37; DB 1; Length 20;  
 Best Local Similarity 77.8%; Pred. No. 3.2;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 DB 8 KKKKKRFSF 16

RESULT 9  
 W1666  
 ID W1666 standard; Peptide; 25 AA.  
 AC W1666;  
 DT 11-JAN-1999 (first entry)  
 DE MARCKs-derived target peptide.  
 KW Calmodulin; green fluorescent protein; GFP; cameleon;  
 KW fluorescence resonance energy transfer; FRET; calcium; sensor;  
 KW analysis; assay; myristoylated alaminte-rich C kinase substrate;  
 KW MARCKs.  
 OS Synthetic.  
 PN WO9840477-A1.  
 PD 17-SEP-1998.  
 PF 13-MAR-1998; U04978.  
 PR 27-AUG-1997; US-919143.  
 PR 14-MAR-1997; US-818252.  
 PR 14-MAR-1997; US-818253.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Miyawaki A, Tsien RY;  
 DR WPI; 98-520809/44.  
 PT New fluorescent protein sensors for detection of analytes -  
 PT comprises a binding protein moiety having an analyte binding region  
 PT and bound donor and acceptor fluorescent protein moieties  
 PS Disclosure; Page 21; 108pp; English.  
 CC This peptide represents a target moiety from myristoylated  
 CC alaminte-rich C kinase substrate (MARCKs) that is recognised by  
 CC calmodulin. The invention provides fluorescent indicators and  
 CC methods for using them to determine the concentration of an  
 CC analyte, such as calcium ion, in vitro and in vivo. Fluorescent  
 CC indicators include a binding protein moiety (e.g. calmodulin) and  
 CC donor and acceptor fluorescent protein moieties, preferably derived  
 CC from Aequorea green fluorescent protein (see W1645-48). The  
 CC binding protein preferably binds target peptides (see W1649-79) in  
 CC addition to the analyte. The target peptide moieties can be  
 CC modified to enhance the response of the fluorescent indicator to  
 CC the analyte.  
 SQ Sequence 25 AA;

Query Match 74.0%; Score 37; DB 1; Length 25;  
 Best Local Similarity 77.8%; Pred. No. 3.9;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 9  
 DB 1 KKKKKRFSF 9

RESULT 10  
 W86744  
 ID W86744 standard; peptide; 30 AA.  
 AC W86744;  
 DT 26-MAR-1999 (first entry)  
 DE Anticoagulant peptide.  
 KW Anticoagulant; blood coagulation inhibitor; disulphide bond; catheter;  
 KW blood bag; dialysis membrane; artificial blood vessel.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Disulfide\_bond 9..24  
 FT Disulfide\_bond 11..22  
 PN J11001493-A.  
 PD 06-JAN-1999.  
 PR 12-JUN-1997; 172827.  
 PR 12-JUN-1997; JP-172827.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI; 99-125476/11.  
 PT New peptide which inhibits blood coagulation - useful in a

PT pharmaceutical material used as a catheter, blood vessel and blood  
 PT dialysis membrane  
 PS Disclosure; Page 6; 18pp; Japanese.  
 CC New inter- or intra-disulphide bonded peptides are disclosed which have  
 CC the formula A-X-Cys(1)-Y-Cys(2)-Z-B A'-X'-Cys(3)-Y'-Cys(4)-Z'-B'  
 CC in which: Cys(1) is disulphide-bonded to Cys(4); Cys(2) is disulphide-  
 CC bonded to Cys(3); A = H or forms a single bond together with B'; B = OH  
 CC or amino or forms a single bond together with A'; A' = H or forms a  
 CC single bond together with B; B' = OH or amino or forms a single bond  
 CC together with A; X and X' = peptide fragments composed of 3 to 13  
 CC amino acid residues; Y and Y' = neutral or basic amino acid residues;  
 CC and Z and Z' = peptide fragments composed of 2 to 12 amino acid  
 CC residues. These peptides inhibit blood coagulation. They can be  
 CC immobilised on the blood-contacting surfaces of catheters, blood  
 CC circuits, blood bags, blood dialysis membranes, artificial blood  
 CC vessels, etc. The present sequence represents a specific example of  
 CC the new peptides.  
 SQ Sequence 30 AA;

Query Match 74.0%; Score 37; DB 1; Length 30;  
 Best Local Similarity 87.5%; Pred. No. 4.6;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 8  
 DB 14 KKKKKKYL 21

RESULT 11  
 W24440  
 ID W24440 standard; peptide; 13 AA.  
 AC W24440;  
 DT 30-SEP-1997 (first entry)  
 DE Nucleic acid (NA) binding peptide used in NA delivery to cells.  
 KW Nucleic acid transporter; gene therapy; binding complex; lysis agent;  
 KW JTS-1; K8; alpha helix; endosome; lysosome; nucleus targeting.  
 OS Synthetic.  
 PN WO9640958-A1.  
 PD 19-DEC-1996.  
 PF 23-APR-1996; U05679.  
 PF 07-JUN-1995; US-484777.  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 PI Smith LC, Sparrow JT, Woo SL;  
 DR WPI; 97-052345/05.  
 PT Nucleic acid transporter useful in gene therapy - contains binding  
 PT complex associated with surface and nuclear ligands and lysis agent  
 PS Disclosure; Page 49; 125pp; English.  
 CC W24434-W24459 are nucleic acid (NA) binding peptides, capable of both  
 CC condensing and stabilising a NA. The peptides can be conjugated to a  
 CC lytic peptide to form a nucleic acid transporter system. The lysis agent  
 CC forms an alpha-helical structure. The transporter system is used to  
 CC deliver nucleic acid to a cell and for treating humans by gene therapy.  
 CC By taking advantage of the characteristics of both the lysis agents  
 CC and the binding molecules, delivery of the nucleic acid is enhanced.  
 CC Specific lysis agents are capable of releasing the nucleic acid  
 CC into the cellular interior from the endosome. Release is efficient  
 CC without endosomal/lysosomal degradation. Once released the binding  
 CC complexes help target the nucleic acid to the nucleus.  
 SQ Sequence 13 AA;

Query Match 72.0%; Score 36; DB 1; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKKKFLF 7  
 DB 6 KKKKKKFLF 12

RESULT 12  
 W22828



ID W22828 standard; peptide; 28 AA.  
AC W22828;  
DT 13-MAR-1998 (first entry)  
DE MICFTR based channel-forming peptide 26.  
KW Channel-forming peptide; channel assembly; epithelial cell; treatment;  
KW cystic fibrosis; polycystic kidney disease; anion transportation; MICFTR.  
OS Synthetic.  
OS Homo sapiens.  
PN WO9726905-A1.  
PD 31-JUL-1997.  
PE 27-JAN-1997; U01103  
PR 24-JAN-1997; US-789155.  
PR 25-JAN-1996; US-591381.  
PR 23-JAN-1997; US-591381.  
PA (UNIV ) UNIV KANSAS MEDICAL CENT.  
PA (UNIV ) UNIV KANSAS STATE RES FOUND.  
PI Iwamoto T, Sullivan LP, Tomich JM;  
DR WPI; 97-393366/36.  
PT Channel assembly for transporting ions across epithelial cell  
PT membranes - comprises new water soluble peptide(s), for treating  
PT cystic fibrosis and polycystic kidney disease by altering water flux  
PT across cells  
PT Example 13: Page 64; 93pp; English.  
PS This MICFTR based peptide is comprised of hydrophobic residues typical of  
CC a transmembrane sequence. The MICFTR peptide associates with membranes  
CC but does not form an ionic conducting channel. This peptide, however with  
CC additional lysine residues at the N-terminal, can be used to construct a  
CC novel channel assembly comprising 3-6 novel peptides, of 18-30 amino  
CC acids. Suitable protein fragments for use as such peptides are present in  
CC the strychnine-binding alpha-subunit of the inhibitory glycine receptor  
CC (M2GlyR) from human brain, and the inhibitory gamma-aminobutyric acid  
CC receptor from human brain, and the cystic fibrosis transmembrane  
CC conductance regulator from human epithelium. The peptides are synthesized  
CC by standard solid phase peptide synthesis. The M2GlyR based peptides are  
CC water soluble to at least 10 mM and enables anions to be transported  
CC through a membrane of an epithelial cell when they are embedded in the  
CC membrane. The channel assembly can be used to alter the flux of water  
CC across an epithelial cell, particularly for treatment of cystic fibrosis  
CC (where affected cells are in the airway, pancreatic duct or epididymis).  
CC The channel assembly can also be used in the treatment of autosomal  
CC dominant polycystic kidney disease (where the affected cells are in the  
CC cystic epithelium). The channel assembly spontaneously inserts into the  
CC basolateral membrane to prevent water flow to adjacent cysts.  
SQ Sequence 28 AA;  
Query Match 72.0%; Score 36; DB 1; Length 28;  
Best Local Similarity 75.0%; Pred. NO. 6.2;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KKKKKFLF 9  
DB 1 KKKKRFMF 8  
RESULT 13  
R05502  
ID R05502 standard; protein; 21 AA.  
AC R05502;  
DT 23-OCT-1990 (first entry)  
DE Modified signal sequence of plasmid pING454.  
KW Thaumatin; sweetener; sugar; sucrose; ds.  
OS Thaumatooccus daniellii.  
PN WO9005775-A.  
PD 31-MAY-1990.  
PE 6-NOV-1989; 005018.  
PR 8-NOV-1988; US-268702.  
PA (ITGE-) Int Genetics Eng Inc.  
PI Blair LC, Koduri RK, Lee JH, Weickmann JI;  
DR WPI; 90-193426/25.  
DR N-PSDB; Q04920.  
PT DNA encoding thaumatin 1 analogues with improved taste -  
PT for use as sweeteners and tools for investigating the taste

PT response.  
PS Example 5; Table 5; 65pp; English.  
CC Thaumatin is 5000 times sweeter in taste than sucrose, and has  
CC potential as sugar substitute. The plasmid contains promoter, an  
CC secretion signal and terminator sequences, and may be used in an  
CC E.coli or S.cerevisiae expression system.  
CC The modified signal sequence increases yield and decreases lag  
CC time for the system.  
SQ Sequence 21 AA;  
Query Match 70.0%; Score 35; DB 1; Length 21;  
Best Local Similarity 77.8%; Pred. NO. 6.9;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KKKKKFLF 10  
DB 3 KKKQAFLEL 11  
RESULT 14  
W12661  
ID W12661 standard; Protein; 584 AA.  
AC W12661;  
DT 13-NOV-1997 (first entry)  
DE Orange pectin methyltransferase (S isoform).  
KW Pectin methyltransferase; PME; orange; de-esterified pectin;  
KW esterification; fruit juice; yoghurt; whey protein; casein;  
KW protein stabilisation.  
OS Citrus sinensis.  
FH Key Location/Qualifiers  
FT peptide 1..46  
FT protein /label= Sig\_peptide  
FT /label= Mat\_protein  
PN WO9703574-A1.  
PD 06-FEB-1997; E03051.  
PR 12-JUL-1996; GB-014438.  
PR 14-JUL-1995; GB-014438.  
PA (DANI-) DANISCO AS.  
PI Buchholt PC, Christensen TMIE, Kreiberg JD, Nielsen J;  
PI Rasmussen H, Thorsoe H;  
DR WPI; 97-145190/13.  
DR N-PSDB; T51739.  
PT Stabilising protein in acidic soln. - by adding enzymatically  
PT de-esterified pectin, produced using pectin methyltransferase  
PS Claim 18; Page 73; 114pp; English.  
CC A polypeptide comprises a novel orange pectin methyltransferase (PME)  
CC that can be obtained by expression of the PME coding sequence  
CC contained in NCIMB 40750 or of a cDNA clone (T51739) obtained from  
CC an orange fruit cDNA library. A longer isoform (see W12662) of the  
CC PME was also identified. Recombinant PME can be expressed in  
CC transformed host cells or transgenic organisms, and used in a  
CC claimed method for de-esterifying a pectin. The de-esterified  
CC pectin is used in claimed methods for stabilising a protein in an  
CC acidic environment (such as fruit juice, drinking yoghurt or drinks  
CC containing whey or milk proteins) without adversely affecting the  
CC viscosity of that environment. The recombinant enzyme can also be  
CC used to esterify pectins in non-aqueous medium. Another novel PME  
CC (see W12660) is also claimed.  
SQ Sequence 584 AA;  
Query Match 70.0%; Score 35; DB 1; Length 584;  
Best Local Similarity 80.0%; Pred. NO. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KKKKKFLF 10  
DB 25 KKKKKFLAL 34  
RESULT 15

W71667  
 ID W71667 standard; Peptide; 24 AA.  
 AC W71667;  
 DT 11-JAN-1999 (first entry)  
 DE MacMARCKS-derived target peptide.  
 KW Calmodulin; green fluorescent protein; GFP; camaleon;  
 KW fluorescence resonance energy transfer; FRET; calcium; sensor;  
 KW analysis; assay; myristoylated alaminte-rich C kinase substrate;  
 KW MacMARCKS.  
 OS Synthetic.  
 PN WO9840477-A1.  
 PD 17-SEP-1998.  
 PF 13-MAR-1998; U04978.  
 PR 27-AUG-1997; US-919143.  
 PR 14-MAR-1997; US-818252.  
 PR 14-MAR-1997; US-818253.  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Miyawaki A, Tsien RY;  
 DR WPI; 98-520809/44.  
 PT New fluorescent protein sensors for detection of analytes -  
 PT comprises a binding protein moiety having an analyte binding region  
 PT and bound donor and acceptor fluorescent protein moieties  
 PS Disclosure; Page 21; 108pp; English.  
 CC This peptide represents a target moiety from myristoylated  
 CC alaminte-rich C kinase substrate (F52 or MacMARCKS) recognised  
 CC by calmodulin. The invention provides fluorescent indicators and  
 CC methods for using them to determine the concentration of an  
 CC analyte, such as calcium ion, in vitro and in vivo. Fluorescent  
 CC indicators include a binding protein moiety (e.g. calmodulin) and  
 CC donor and acceptor fluorescent protein moieties, preferably derived  
 CC from Aequorea green fluorescent protein (see W71645-48). The  
 CC binding protein preferably binds target peptides (see W71649-79) in  
 CC addition to the analyte. The target peptide moieties can be  
 CC modified to enhance the response of the fluorescent indicator to  
 CC the analyte.  
 SQ Sequence 24 AA;

Query Match 70.0%; Score 35; DB 1; Length 24;  
 Best Local Similarity 87.5%; Pred. No. 7.7;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KKKKKELF 9  
 |||||  
 Db 1 KKKKKFSF 8

Search completed: September 7, 1999, 23:18:21  
 Job time: 1677 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:17:32 ; Search time 41.46 Seconds  
(without alignments)  
14.844 Million cell updates/sec

Title: US-09-124-280A-43

Perfect score: 62

Sequence: 1 CKFKFKFKFC 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

- SPTREMBL10:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	62.9	212	2	086966 desulfovibr
2	38	61.3	275	2	033966 desulfovibr
3	38	61.3	208	2	086964 desulfovibr
4	38	61.3	223	2	086970 desulfovibr
5	38	61.3	367	2	086974 desulfovibr
6	38	61.3	213	2	086976 desulfovibr
7	38	61.3	221	2	086978 desulfovibr
8	38	61.3	471	2	092818 desulfovibr
9	38	61.3	121	12	Q9P01 chorisoneu
10	37	59.7	190	2	024866 helicobacte
11	37	59.7	191	2	Q92N40 helicobacte
12	37	59.7	182	11	P97721 rattus norv
13	36	58.1	175	5	060943 oxytricha f
14	36	58.1	622	5	044756 caenorhabdi
15	36	58.1	224	5	Q20115 caenorhabdi
16	36	58.1	273	10	Q23527 arabidopsis
17	36	58.1	379	11	Q88852 mus musculu
18	36	58.1	679	13	Q98856 cynops pyr
19	35	56.5	434	3	Q02933 saccharomyc
20	35	56.5	433	4	Q94828 homo sapien
21	35	56.5	508	5	Q93227 caenorhabdi
22	35	56.5	779	10	Q81287 arabidopsis
23	34	54.8	170	1	Q59556 pyrococcus
24	34	54.8	61	2	Q47617 escherichia
25	34	54.8	449	2	Q92M35 helicobacte
26	34	54.8	504	2	Q92F22 bacillus st
27	34	54.8	560	4	Q14520 homo sapien
28	34	54.8	650	5	Q44236 clona lites
29	34	54.8	3345	5	Q46074 drosophila

30	34	54.8	536	5	Q19294 caenorhabdi
31	34	54.8	946	5	Q62080 caenorhabdi
32	34	54.8	439	10	O81424 arabidopsis
33	34	54.8	422	10	O04260 arabidopsis
34	34	54.8	662	11	P97581 rattus norv
35	34	54.8	246	12	Q9Y87 melanolus
36	34	54.8	59	12	Q9YR13 porcine cir
37	34	54.8	59	12	Q9YQTO porcine cir
38	33	53.2	397	1	Q33909 pyrobaculum
39	33	53.2	260	2	Q53852 spiroplasma
40	33	53.2	322	2	Q53848 spiroplasma
41	33	53.2	1125	3	Q07660 saccharomyc
42	33	53.2	893	4	Q94976 homo sapien
43	33	53.2	606	5	O44940 drosophila
44	33	53.2	561	5	Q17731 caenorhabdi
45	33	53.2	606	5	O46104 drosophila

ALIGNMENTS

RESULT 1  
O86966  
ID O86966 PRELIMINARY; PRT; 212 AA.  
AC O86966;  
DT 01-NOV-1998 (TEMBLrel. 08, Created)  
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
OS Desulfovibrio sp. PIB2.  
OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN-PIB2.  
RX MEDLINE; 98269016.  
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
RT "Phylogeny of dissimilatory sulfite reductases supports an early  
RT origin of sulfate respiration."  
RL J. Bacteriol. 180:2975-2982(1998).  
DR EMBL: U58116; AAC24107.1; -;  
FT NON\_TER 1  
FT NON\_TER 212  
SQ SEQUENCE 212 AA; 23441 MW; C90DA46D CRC32;

Query Match 62.9%; Score 39; DB 2; Length 212;  
Best Local Similarity 77.8%; Pred. No. 7.5;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKFC 10  
DB 145 KFKFKFDAC 153

RESULT 2  
O33966  
ID O33966 PRELIMINARY; PRT; 275 AA.  
AC O33966;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)  
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (EC 1.8.99.1)  
DE (FRAGMENT).  
GN DSR.  
OS Desulfovibrio simplex.  
OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN-XVI (DSM4141);  
RA POTT W.M., THUM-SCHMITZ A.S., FAATH I., DAHL C., TRUEPER H.G.;  
RL Microbiology 0:0-0(1997).  
RN [2]  
RP SEQUENCE FROM N.A.

```
RC STRAIN-XVI (DSM4141);
RA MOLITOR M., DAHL C., ZELLNER G., TRUEPER H.G.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: H(2)S + ACCEPTOR + 3 H(2)O = SULFITE +
CC REDUCED ACCEPTOR.
CC -!- COFACTOR: IRON.
DR EMBL; U78738; AAB66716.1; -.
KW Oxidoreductase.
FT NON_TER 275
SQ SEQUENCE 275 AA; 30612 MW; F7EC82B5 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 275;
Best Local Similarity 77.8%; Pred. No. 14; 2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

QY 2 KFKFKFKFC 10
DB 213 KFKFKFDGC 221

RESULT 3
ID O86964 PRELIMINARY; PRT; 208 AA.
AC O86964;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).
OS Desulfovibrio sp. PT-2.
OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-PT-2;
RX MEDLINE; 98269016.
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;
RT "Phylogeny of dissimilatory sulfite reductases supports an early
RT origin of sulfate respiration."
RL J. Bacteriol. 180:2975-2982(1998).
DR EMBL; U58114; AAC24105.1; -.
FT NON_TER 1
FT NON_TER 208
SQ SEQUENCE 208 AA; 23141 MW; 79E7999F CRC32;

Query Match 61.3%; Score 38; DB 2; Length 208;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKFC 10
DB 145 KFKFKFDGC 153

RESULT 4
ID O86970 PRELIMINARY; PRT; 223 AA.
AC O86970;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).
OS Desulfohalobium sapovorans.
OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 33892;
RX MEDLINE; 98269016.
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;
RT "Phylogeny of dissimilatory sulfite reductases supports an early
RT origin of sulfate respiration."
RL J. Bacteriol. 180:2975-2982(1998).
DR EMBL; U58120; AAC24109.1; -.

```

```
FT NON_TER 1
FT NON_TER 223
SQ SEQUENCE 223 AA; 24764 MW; FD37B9CC CRC32;

Query Match 61.3%; Score 38; DB 2; Length 223;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKFC 10
DB 145 KFKFKFDGC 153

RESULT 5
ID O86974 PRELIMINARY; PRT; 367 AA.
AC O86974;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).
OS Desulfohalobium latius.
OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 43918;
RX MEDLINE; 98269016.
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;
RT "Phylogeny of dissimilatory sulfite reductases supports an early
RT origin of sulfate respiration."
RL J. Bacteriol. 180:2975-2982(1998).
DR EMBL; U58124; AAC24097.1; -.
FT NON_TER 1
FT NON_TER 367
SQ SEQUENCE 367 AA; 41543 MW; AACB5A60 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 367;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKFC 10
DB 145 KFKFKFDGC 153

RESULT 6
ID O86976 PRELIMINARY; PRT; 213 AA.
AC O86976;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).
OS Desulfohalobium multivorans.
OC Bacteria; Proteobacteria; delta subdivision; Desulfohalobium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 33890;
RX MEDLINE; 98269016.
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;
RT "Phylogeny of dissimilatory sulfite reductases supports an early
RT origin of sulfate respiration."
RL J. Bacteriol. 180:2975-2982(1998).
DR EMBL; U58126; AAC24101.1; -.
FT NON_TER 1
FT NON_TER 213
SQ SEQUENCE 213 AA; 23637 MW; F50D1C25 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 213;
Best Local Similarity 77.8%; Pred. No. 11;

```

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFKFKFKFC 10  
| | | | |  
Db 145 KFKFKFDGC 153

## RESULT 7

OS 086978 PRELIMINARY; PRT; 221 AA.  
AC 086978;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
DE DISSIMILATORY SULFITE REDUCTASE ALPHA SUBUNIT (FRAGMENT).  
OS Desulfonema limicola.  
OC Bacteria; Proteobacteria; delta subdivision; Desulfonema.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 33961;  
RX MEDLINE; 98269016.  
RA WAGNER M., ROGER A.J., FLAX J.L., BRUSSEAU G.A., STAHL D.A.;  
RT "Phylogeny of dissimilatory sulfite reductases supports an early  
RT origin of sulfate respiration."  
RL J. Bacteriol. 180:2975-2982(1998).  
DR EMBL; U58128; AAC24099.1; -.  
FT NON\_TER 1  
FT NON\_TER 221  
FT NON\_TER 221  
SO SEQUENCE 221 AA; 24396 MW; C81B1A78 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 221;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFKFKFKFC 10  
| | | | |  
Db 145 KFKFKFDGC 153

## RESULT 8

OS 092H18 PRELIMINARY; PRT; 471 AA.  
AC 092H18;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DE DISSIMILATORY SULFITE REDUCTASE SUBUNIT A.  
GN DSR4.  
OS Desulfotomaculum thermocisternum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Desulfotomaculum.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ST90;  
RA LARSEN O., LIEN T., BIRKELAND N.-K.;  
RT "Dissimilatory sulfite reductase from Archaeoglobus profundus and  
RT Desulfotomaculum thermocisternum: phylogenetic and structural  
RT implications from gene sequences."  
RL Extremophiles 0:0-0(1999).  
DR EMBL; AF074396; AAC96107.1; -.  
SO SEQUENCE 471 AA; 54100 MW; 54C2B846 CRC32;

Query Match 61.3%; Score 38; DB 2; Length 471;  
Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFKFKFKFC 10  
| | | | |  
Db 247 KFKFKFDGC 255

## RESULT 9

Q9YP01  
AC Q9YP01;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DE 15 KDA PROTEIN HOMOLOG.  
OS Choristoneura fumiferana nuclear polyhedrosis virus (CfMNPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-IRELAND;  
RA POLOUMIENKO A., KRELL P.J.;  
RT "Identification of the GTA (probable global transactivator) gene of a  
RT baculovirus pathogenic to the spruce budworm, CfMNPV."  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U72240; AAD10315.1; -.  
SO SEQUENCE 121 AA; 13337 MW; E2D4D213 CRC32;

Query Match 61.3%; Score 38; DB 12; Length 121;  
Best Local Similarity 86.7%; Pred. No. 7;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFKFKFKF 9  
| | | | |  
Db 111 CKLAKLKF 119

## RESULT 10

OS 024866 PRELIMINARY; PRT; 190 AA.  
AC 024866;  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DE HYPOTHETICAL 21.2 KD PROTEIN.  
GN HP0021.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-26695;  
RX MEDLINE; 97394467.  
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,  
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,  
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,  
RA VENTER J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori."  
RL Nature 388:539-547(1997).  
DR EMBL; AE000524; AAD07094.1; -.  
DR TIGR; HP0021; -.  
KW Hypothetical protein.  
SO SEQUENCE 190 AA; 21183 MW; 824A70B9 CRC32;

Query Match 59.7%; Score 37; DB 2; Length 190;  
Best Local Similarity 77.8%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFKFKFKFC 10  
| | | | |  
Db 3 KFLFKQKFC 11

```

RESULT 11
Q92N40 PRELIMINARY; PRT; 191 AA.
AC Q92N40;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE PUTATIVE.
GN JHP0019.
OS Helicobacter pylori J99.
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RC STRAIN=J99;
RA ALM R.A., LING L.-S.L., MOIR D.T., KING B.L., BROWN E.D., DOIG P.C.,
RA SMITH D.R., NOONAN B., GUILD B.C., DEJONGE B.L., CARMEL G.,
RA TUMMINO P.J., CARUSO A., URIA-NICKELSEN M., MILLS D.M., IVES C.,
RA GIBSON R., MERBERG D., MILLS S.D., JIANG Q., TAYLOR D.E., VOVIS G.F.,
RA TRUST T.J.;
RT "Genomic Sequence Comparison of Two Unrelated Isolates of the Human
RT Gastric Pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001441; AAD05601.1; -.
SQ SEQUENCE 191 AA; 21151 MW; 28579193 CRC32;

Query Match 59.7%; Score 37; DB 2; Length 191;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKC 10
DB 3 KLFKQKFC 11

RESULT 12
P97721 PRELIMINARY; PRT; 182 AA.
AC P97721;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE RBSC2-NA-K-CL COTRANSPORTER HOMOLOG (FRAGMENT).
GN RBSC2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE; 96184669.
RA HIDAKA H., OSHIMA T., IKEDA K., FUKUKAWA M., TAKASAKA T.;
RT "The Na-K-Cl cotransporters in the rat cochlea: RT-PCR and partial
RT sequence analysis.";
RL Biochem. Biophys. Res. Commun. 220:425-430(1996).
DR EMBL; S82233; CAB34010.1; -.
FT NON_TER 1
SQ SEQUENCE 182 AA; 19189 MW; 8281FB5A CRC32;

Query Match 59.7%; Score 37; DB 11; Length 182;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKFKFKFKC 10
DB 68 CKLNFDFSYC 77

RESULT 13
O60943 PRELIMINARY; PRT; 175 AA.
ID O60943
AC O60943;

```

```

DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE 20 KDA PROTEIN.
OS Oxytricha fallax.
OC Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;
OC Oxytrichidae; Oxytricha.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN-SUBKARYONIDE 3.5;
RX MEDLINE; 86079559.
RA HERRICK G., CARTINHOOR S., DAWSON D., ANG D., SHEETS R., LEE A.,
RA WILLIAMS K.;
RT "Mobile elements bounded by C4A4 telomeric repeats in Oxytricha
RT fallax.";
RL Cell 43:759-768(1985).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN-SUBKARYONIDE 3.5;
RX MEDLINE; 90185189.
RA HUNTER D.J., WILLIAMS K., CARTINHOOR S., HERRICK G.;
RT "Precise excision of telomere-bearing transposons during Oxytricha
RT fallax macronuclear development.";
RL Genes Dev. 3:2101-2112(1989).
RN [3]
RC SEQUENCE FROM N.A.
RC STRAIN-SUBKARYONIDE 3.5;
RA DOAK T.G., WILLIAMS K., HERRICK G.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE FROM N.A.
RC STRAIN-SUBKARYONIDE 3.5;
RX MEDLINE; 94134747.
RA DOAK T.G., DOERDER F.P., JAHN C.L., HERRICK G.;
RT "A proposed superfamily of transposase genes: transposon-like
RT elements in ciliated protozoa and a common "D35E" motif.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).
DR EMBL; L23169; AAA18576.1; -.
SQ SEQUENCE 175 AA; 19981 MW; F062C7E6 CRC32;

Query Match 58.1%; Score 36; DB 5; Length 175;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KFKFKFK 9
DB 143 KFKFKFY 149

RESULT 14
O44756 PRELIMINARY; PRT; 622 AA.
ID O44756
AC O44756;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE R13H8.1 PROTEIN.
GN R13H8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN-BRISTOL R2;
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPESE T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

```

RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]

RN SEQUENCE FROM N.A.  
 RP STRAIN-BRISTOL N2;  
 RC JONES K., HINDS K., SUTTERER C., COPMAN M.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RN SEQUENCE FROM N.A.  
 RP STRAIN-BRISTOL N2;  
 RC WATERSTON R.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF039717: AAB96741.1: "-".  
 DR PFAM: PF00250; Fork\_head; 1.

SQ SEQUENCE 622 AA; 69110 MW; 1F459261 CRC32;

Query Match 58.1%; Score 36; DB 5; Length 622;  
 Best Local Similarity 66.7%; Pred. No. 57;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 KFKFKFKC 10  
 I |||||  
 Db 574 KFKFKFKC 582

## RESULT 15

Q20115 PRELIMINARY; PRT; 224 AA.  
 AC Q20115; Q21912;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE R10E9.3 PROTEIN.  
 GN R10E9.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GARDNER A.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z35600; CAA84665.1; "-".  
 DR EMBL: Z35601; CAA84665.1; JOINED.  
 DR EMBL: Z35601; CAA84668.1; "-".  
 DR EMBL: Z35600; CAA84668.1; JOINED.  
 SQ SEQUENCE 224 AA; 26004 MW; 8D5E3815 CRC32;

Query Match 58.1%; Score 36; DB 5; Length 224;  
 Best Local Similarity 75.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CKFKFKFK 8  
 I |||||  
 Db 84 CNFKFKLK 91

Search completed: September 7, 1999, 23:17:34  
 Job time: 1777 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 1999, 00:43:00 ; Search time 21.13 Seconds  
(without alignments)  
13.378 Million cell updates/sec

Title: US-09-124-280A-43  
Perfect score: 62  
Sequence: 1 CKFKFKFKC 10  
Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues  
Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	66.1	586	1	P33919 escherichia
2	39	62.9	436	1	P45574 desulfovibr
3	37	59.7	663	1	Q90611 gallus gall
4	37	59.7	1205	1	P50112 mus musculu
5	36	58.1	350	1	P31561 euglena gra
6	35	56.5	434	1	P53916 saccharomyc
7	34	54.8	660	1	P08253 homo sapien
8	34	54.8	662	1	P33434 mus musculu
9	34	54.8	662	1	P50757 oryctolagus
10	34	54.8	662	1	P33436 rattus norv
11	34	54.8	61	1	P03854 escherichia
12	34	54.8	61	1	P13946 escherichia
13	34	54.8	449	1	P56125 helicobacte
14	34	54.8	184	1	P80691 carica papa
15	34	54.8	2483	1	P18490 drosophila
16	34	54.8	453	1	Q11095 caenorhabdi
17	33	53.2	355	1	P39857 staphylococ
18	33	53.2	846	1	P11584 drosophila
19	33	53.2	430	1	P44689 haemophilus
20	33	53.2	145	1	P31093 hordeum vul
21	33	53.2	66	1	P80950 androctonus
22	33	53.2	593	1	Q80922 human papil
23	33	53.2	322	1	P15894 spiroplasma
24	33	53.2	364	1	P22727 mus musculu
25	33	53.2	201	1	P37017 escherichia
26	32.5	52.4	104	1	P22069 rana pipien
27	32	51.6	1290	1	P10844 clostridium
28	32	51.6	417	1	Q33998 chromatium
29	32	51.6	133	1	P81017 echis carin
30	32	51.6	1212	1	P55011 homo sapien
31	32	51.6	1121	1	Q39557 ceratodon p
32	32	51.6	203	1	P43487 homo sapien
33	32	51.6	201	1	P34022 mus musculu
34	32	51.6	337	1	P46756 acanthamoeb
35	32	51.6	2470	1	P35169 saccharomyc
36	32	51.6	381	1	Q86917 capripoxvir
37	32	51.6	178	1	P47561 mycoplasma
38	32	51.6	1361	1	Q04693 saccharomyc
39	31	50.0	3305	1	Q25490 manduca sex
40	31	50.0	221	1	P32765 theobroma c
41	31	50.0	504	1	P22680 homo sapien
42	31	50.0	501	1	P51542 oryctolagus
43	31	50.0	459	1	P53452 fuigu rubrip

ALIGNMENTS

```
RESULT 1
YEJH_ECOLI
ID YEJH_ECOLI STANDARD; PRT; 586 AA.
AC P33919; P36926; P36927; P76449;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 66.4 KD PROTEIN IN RSUA-REPLY INTERGENIC REGION.
GN YEJH.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / BHB2600;
RA RICHTERICH P., LAKEY N., GRYAN G., JAEHN L., MINTZ L., ROBISON K.,
RA CHURCH G.M.;
RL SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN RP
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
SCIENCE 277:1433-1474(1997).
CC -1- SIMILARITY: SOME SIMILARITY TO A PHAGE PROTEIN AND RESTRICTION-
MODIFICATION SYSTEMS.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
THAT PRODUCE THREE SEPARATE ORFS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; U00008; G405911; ALT_FRAME.
DR EMBL; AE000308; G1788511; -.
DR ECOGENE; EG12045; YEJH.
DR PFAM; PF00271; helicase_C; 1.
RW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 586 AA; 66413 MW; EE8BA2E1 CRC32;
```

Query Match 66.1%; Score 41; DB 1; Length 586;  
Best Local Similarity 60.0%; Pred. No. 3.3;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFKFKFKC 10  
| : : : : :  
Db 425 CDFRFRKNC 434

```
RESULT 2
DSVA_DESVH
ID DSVH_DESVH STANDARD; PRT; 436 AA.
AC P45574; Q46581;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SULFITE REDUCTASE, DISSIMILATORY-TYPE ALPHA SUBUNIT (EC 1.8.99.3)
```

DE (DESULFOVIRIDIN ALPHA SUBUNIT) (HYDROGENSULFITE REDUCTASE ALPHA  
DE SUBUNIT).

GN DSVN.

OS DESULFOVIRIDIN VULGARIS (STRAIN HILDENBOROUGH).

OC BACTERIA; PROTEOBACTERIA; DELTA SUBDIVISION; DESULFOVIRIDIO.

RN [1].

RP SEQUENCE FROM N.A.

RX MEDLINE; 9519398.

RA KARKHOFF-SCHWEIZER R.R., HUBER D.P.W., VOORDOUN G.;  
RT "Conservation of the genes for dissimilatory sulfite reductase from  
RT Desulfovibrio vulgaris and Archaeoglobus fulgidus allows their  
RT detection by PCR."

RL APPL. ENVIRON. MICROBIOL. 61:290-296(1995).

RN [2].

RP SEQUENCE OF 1-25.

RX MEDLINE; 92209491.

RA PIERIK A.J., DUYVIS M.G., VAN HELVOORT J.M.L.M., WOLBERT R.B.G.,  
RA HAGEN W.R.;  
RT "The third subunit of desulfovibridin-type dissimilatory sulfite  
RT reductases."

RL EUR. J. BIOCHEM. 205:111-115(1992).

CC -!- FUNCTION: CATALYZES THE REDUCTION OF SULFITE TO SULFIDE. THIS IS  
CC THE TERMINAL OXIDATION REACTION IN SULFATE RESPIRATION, A PROCESS  
CC CATALYZED BY THE SULFATE-REDUCING BACTERIA.

CC -!- CATALYTIC ACTIVITY: O(3S.S.SO(3))(2-) + ACCEPTOR + 2 H(2)O +  
CC OH(-) = 3 HSO(-)(3) + REDUCED ACCEPTOR.

CC -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA, TWO BETA AND TWO GAMMA  
CC SUBUNITS.

CC -!- COFACTOR: IRON-SULFUR CLUSTERS AND TWO SIROHEMES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U16723; G902747; -  
DR PIR; S21197; S21197.  
KW OXIDOREDUCTASE; HEME; IRON-SULFUR; 4FE-4S.  
FT INIT\_MET 0  
FT METAL 176 HEME IRON (POTENTIAL).  
FT METAL 182 HEME IRON (POTENTIAL).  
FT METAL 220 HEME IRON (POTENTIAL).  
FT METAL 224 HEME IRON (POTENTIAL).  
FT METAL 283 IRON-SULFUR (4FE-4S) (POTENTIAL).  
FT METAL 302 IRON-SULFUR (4FE-4S) (POTENTIAL).  
FT METAL 305 IRON-SULFUR (4FE-4S) (POTENTIAL).  
FT METAL 308 IRON-SULFUR (4FE-4S) (POTENTIAL).  
FT CONFLICT 21 S -> V (IN REF. 2).  
FT CONFLICT 25 Q -> K (IN REF. 2).  
SQ SEQUENCE 436 AA; 48959 MW; F91ECID6 CRC32;

Query Match 62.9%; Score 39; DB 1; Length 436;  
Best Local Similarity 77.8%; Pred.No. 5.5;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KFKFKFKFC 10  
|||||  
Db 212 KFKFKFDAC 220

RESULT 3  
COG2\_CHICK STANDARD; PRT; 663 AA.  
AC Q90611;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE 72 KD TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KD GELATINASE)  
DE (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).

MMP2.

GN GALLUS GALLUS (CHICKEN).

OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.

RN [1].

RP SEQUENCE FROM N.A.

RX TISSUE-EMBRYO;  
MEDLINE; 94280397.

RA AMES R.T., FRENCH D.L., QUIGLEY J.P.;  
RT "Cloning of a 72 kDa matrix metalloproteinase (gelatinase) from  
RT chicken embryo fibroblasts using gene family PCR: expression of the  
RT gelatinase increases upon malignant transformation."

RL BIOCHEM. J. 300:729-736(1994).

RN [2].

RP SEQUENCE OF 27-41 AND 107-122.

RX MEDLINE; 91161603.

RA CHEN J.-M., AMES R.T., WARD G.R., YOUNGLEIB G.L., QUIGLEY J.P.;  
RT "Isolation and characterization of a 70-kDa metalloprotease  
RT (gelatinase) that is elevated in Rous sarcoma virus-transformed  
RT chicken embryo fibroblasts."

RL J. BIOL. CHEM. 266:5113-5121(1991).

CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLI-|-  
CC ILE-ALA-GLY-GLN.

CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.

CC -!- TISSUE SPECIFICITY: PRODUCED BY NORMAL SKIN FIBROBLASTS.

CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN

CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U07775; G504476; -  
DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
DR PFAM; PF00040; fn2; 3.  
DR PFAM; PF00045; hemopexin; 4.  
DR PFAM; PF00099; zn-protease; 1.  
DR PFAM; PF00413; matrixin; 1.  
DR HSP; P08253; 1RTG.  
KW HYDROLASE; METALLOPROTEASE; ZINC; ZYMOGEN; CALCIUM;  
KW COLLAGEN DEGRADATION; EXTRACELLULAR MATRIX; SIGNAL.  
FT SIGNAL 1 26  
FT PROPEP 27 106 ACTIVATION PEPTIDE.  
FT CHAIN 107 663 72 KD TYPE IV COLLAGENASE.  
FT DOMAIN 97 103 AUTOINHIBITOR REGION.  
FT DOMAIN 107 218 COLLAGENASE-LIKE.  
FT DOMAIN 219 393 COLLAGEN-BINDING.  
FT DOMAIN 394 468 COLLAGENASE-LIKE.  
FT DOMAIN 469 663 HEMOPEXIN-LIKE.  
FT DOMAIN 223 280 FIBRONECTIN TYPE-II.  
FT DOMAIN 281 388 FIBRONECTIN TYPE-II.  
FT DOMAIN 339 396 FIBRONECTIN TYPE-II.  
FT METAL 400 400 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 401 401 BY SIMILARITY.  
FT METAL 404 404 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 410 410 ZINC (CATALYTIC) (BY SIMILARITY).  
FT DISULFID 472 663 BY SIMILARITY.  
FT CONFLICT 40 P -> Q (IN REF. 2).  
FT CONFLICT 116 W -> T (IN REF. 2).  
FT CONFLICT 122 T -> I (IN REF. 2).  
SQ SEQUENCE 663 AA; 74941 MW; 66F15FDC CRC32;

Query Match 59.7%; Score 37; DB 1; Length 663;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKKFKFK 8

Db 288 CKFPFREQ 295

RESULT 4

ID NKCL\_MOUSE STANDARD; PRT; 1205 AA.  
AC P55012;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE BUNETANIDE-SENSITIVE SODIUM-(POTASSIUM)-CHLORIDE COTRANSPORTER 1  
DE (BASOLATERAL NA-K-CL SYMPORTER).  
GN SLC12A2 OR NKCL1.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95014374.  
RA DELPIRE E., RAUCHMAN M.I., BEIER D.R., HEBERT S.C., GULLANS S.R.;  
RT "Molecular cloning and chromosome localization of a putative  
basolateral Na(+)-K(+)-2Cl-cotransporter from mouse inner medullary  
collecting duct (mMCD-3) cells";  
RL J. BIOL. CHEM. 269:25877-25883(1994).  
CC -1- FUNCTION: ELECTRICALLY SILENT TRANSPORTER SYSTEM WHICH IS A  
MEDIATOR OF SODIUM AND CHLORIDE REABSORPTION. PLAYS A VITAL ROLE  
IN THE REGULATION OF IONIC BALANCE AND CELL VOLUME.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE SLC12A FAMILY OF TRANSPORTERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U13174; G560008; -  
DR MGD; MGI:101924; SLC12A2.  
DR PFAM; PF00324; aa\_permeases; 1.  
KW TRANSPORT; TRANSMEMBRANE; GLYCOPROTEIN.  
FT DOMAIN 1 278 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 279 299 POTENTIAL.  
FT DOMAIN 300 302 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 303 323 POTENTIAL.  
FT DOMAIN 324 360 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 361 381 POTENTIAL.  
FT DOMAIN 382 403 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 404 424 POTENTIAL.  
FT DOMAIN 425 428 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 429 449 POTENTIAL.  
FT DOMAIN 450 479 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 480 500 POTENTIAL.  
FT DOMAIN 501 513 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 514 534 POTENTIAL.  
FT DOMAIN 535 584 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 585 605 POTENTIAL.  
FT DOMAIN 606 650 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 651 671 POTENTIAL.  
FT DOMAIN 672 692 POTENTIAL.  
FT TRANSMEM 693 709 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 710 730 POTENTIAL.  
FT TRANSMEM 731 892 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 893 913 POTENTIAL.  
FT DOMAIN 914 1205 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 90 107 POLY-ALA.

FT CARBOHYD 546 546 POTENTIAL.  
FT CARBOHYD 555 555 POTENTIAL.  
SQ SEQUENCE 1205 AA; 130949 MW; 2B79C1BA CRC32;

Query Match 59.7%; Score 37; DB 1; Length 1205;  
Best Local Similarity 50.0%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKKFKFKFC 10

Db 561 CKLNDFESYC 570

RESULT 5

ID YCKA\_EUGGR STANDARD; PRT; 350 AA.  
AC P31561;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 40.9 KD PROTEIN IN 16S RRNA 3'REGION (ORF350).  
OS EUGLENA GRACILIS.  
OG CHLOROPLAST.  
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-2;  
RX MEDLINE; 93347989.  
RA HALLICK R.B., HONG L., DRAGER R.G., FAVREAU M.R., MONFORT A.,  
RA ORSAT B., SPIELMANN A., STUTZ E.;  
RT "Complete sequence of Euglena gracilis chloroplast DNA";  
RL NUCLEIC ACIDS RES. 21:3537-3544(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN-2;  
RA SCHLUNEGGER B., STUTZ E.;  
RT "The Euglena gracilis chloroplast genome: structural features of a  
DNA region possibly carrying the single origin of DNA replication.";  
RL CURR. GENET. 8:629-634(1984).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z11874; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X70810; G415796; -  
DR PIR; S34557; S34557.  
DR PIR; S36933; S36933.  
KW CHLOROPLAST; HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 350 AA; 40861 MW; 6EF72410 CRC32;

Query Match 58.1%; Score 36; DB 1; Length 350;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKKFKFKFC 9

Db 167 CAFNFAKFP 175

RESULT 6

ID YNM8\_YEAST STANDARD; PRT; 434 AA.  
AC P53916;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 50.2 KD PROTEIN IN CPT1-SPC98 INTERGENIC REGION.  
GN YNL129W OR N1220 OR N1872.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCHAROMYCETALES;  
RN SACCHAROMYCETACEAE; SACCHAROMYCES.  
RC [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C.  
RX MEDLINE; 96109932.  
RA MAULLET L., BOUSSEREAU F., JACQUET M.;  
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
RT MEP2, CAP/SRV2, NAM9, FKBI/FPRI/RBP1, MOM22 and CPT1, predicts an  
RT adenosine deaminase gene and 14 new open reading frames.";  
RL YEAST 11:1195-1209(1995).  
CC -1- SIMILARITY: SOME, TO PROTEIN-TYROSINE PHOSPHATASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 246843; G854509; -.  
DR EMBL; 271404; E239804; -.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS00556; TYR\_PHOSPHATASE\_2; 1.  
KW HYPOTHETICAL PROTEIN; HYDROLASE.  
SQ SEQUENCE 434 AA; 50152 MW; 368BEB97 CRC32;  
-----  
Query Match 56.5%; Score 35; DB 1; Length 434;  
Best Local Similarity 85.7%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 2 KEKFKFK 8  
| | | | | | |  
DB 268 KEKFKFR 274  
-----  
RESULT 7  
ID COG2\_HUMAN STANDARD; PRT; 660 AA.  
AC P08253;  
DT 01-AUG-1988 (REL. 08, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE 72 KD TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KD GELATINASE)  
DE (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A) (TBE-1).  
GN MMP2 OR CLG4A.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE OF 19-660 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 98198218.  
RA COLLIER I.E., WILHELM S.M., EISEN A.Z., MARNER B.L., GRANT G.A.,  
RA SELTZER J.L., KRONBERGER A., HE C., BAUER E.A., GOLDBERG G.I.;  
RT "H-ras oncogene-transformed human bronchial epithelial cells (TBE-1)  
RT secrete a single metalloprotease capable of degrading basement  
RT membrane collagen."  
RL J. BIOL. CHEM. 263:6579-6587(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91236162.  
RA COLLIER I.E., BRUNS G.A.P., GOLDBERG G.I., GERHARD D.S.;  
RT "On the structure and chromosome location of the 72- and 92-kDa human  
RT type IV collagenase genes."  
RL GENOMICS 9:429-434(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90293047.

RA HUHTALA P., CHOW L.T., TRYGGVASON K.;  
RT "Structure of the human type IV collagenase gene.";  
RL J. BIOL. CHEM. 265:11077-11082(1990).  
RN [4]  
RP SEQUENCE OF 1-51 FROM N.A.  
RX MEDLINE; 90228972.  
RA HUHTALA P., EDDY R.L., FAN Y.S., BYERS M.G., SHOWS T.B.,  
RA TRYGGVASON K.;  
RT "Completion of the primary structure of the human type IV collagenase  
RT preproenzyme and assignment of the gene (CLG4) to the q21 region of  
RT chromosome 16.";  
RL GENOMICS 6:554-559(1990).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 443-660.  
RX MEDLINE; 96069777.  
RA LIBSON A.M., GITTS A.G., COLLIER I.E., MARNER B.L., GOLDBERG G.I.,  
RA LATTMAN E.E.;  
RT "Crystal structure of the haemopexin-like C-terminal domain of  
RT gelatinase A.";  
RL NAT. STRUCT. BIOL. 2:938-942(1995).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 458-660.  
RX MEDLINE; 96140723.  
RA GOELKE U., GOMIS-RUTH F.X., CRABBE T., MURPHY G., DOCHERTY A.J.,  
RA BODE W.;  
RT "The C-terminal (haemopexin-like) domain structure of human  
RT gelatinase A (MMP2): structural implications for its function.";  
RL FEBS LETT. 378:126-130(1996).  
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-[-  
CC ILE-ALA-GLY-GLN.  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY NORMAL SKIN FIBROBLASTS.  
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J03210; G180671; -.  
DR EMBL; M33789; G180601; -.  
DR EMBL; M55593; G180616; -.  
DR EMBL; M58552; G180616; JOINED.  
DR EMBL; M55582; G180616; JOINED.  
DR EMBL; M55583; G180616; JOINED.  
DR EMBL; M55584; G180616; JOINED.  
DR EMBL; M55585; G180616; JOINED.  
DR EMBL; M55586; G180616; JOINED.  
DR EMBL; M55587; G180616; JOINED.  
DR EMBL; M55588; G180616; JOINED.  
DR EMBL; M55589; G180616; JOINED.  
DR EMBL; M55590; G180616; JOINED.  
DR EMBL; M55591; G180616; JOINED.  
DR EMBL; M55592; G180616; JOINED.  
DR PIR; A28153; A28153.  
DR PDB; 1RTG; 10-JUN-96.  
DR PDB; 1GEN; 17-AUG-96.  
DR MIN; 120360; -.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
DR PFAM; PF00040; fn2; 3.  
DR PFAM; PF00045; hemopexin; 4.  
DR PFAM; PF00099; zn-protease; 1.  
DR PFAM; PF00413; matrixin; 1.

KW HYDROLASE; METALLOPROTEIN; ZINC; ZMOGEN; CALCIUM;  
 KW COLLAGEN DEGRADATION; EXTRACELLULAR MATRIX; DUPLICATION; SIGNAL;  
 3D-STRUCTURE.  
 FT SIGNAL 1 29  
 FT PROPEP 30 109  
 FT CHAIN 110 660  
 FT DOMAIN 100 107  
 FT DOMAIN 110 221  
 FT DOMAIN 222 396  
 FT DOMAIN 397 465  
 FT DOMAIN 466 660  
 FT DOMAIN 226 283  
 FT DOMAIN 284 341  
 FT DOMAIN 342 399  
 FT METAL 403 404  
 FT METAL 407 407  
 FT METAL 413 413  
 FT CARBOHYD 573 573  
 FT DISULFID 642 642  
 FT DISULFID 469 660  
 SQ SEQUENCE 660 AA; 73882 MW; 4372D672 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 660;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CKFKFK 8  
 DB 291 CKFFPRFQ 298

RESULT 8  
 COG2\_MOUSE  
 ID COG2\_MOUSE STANDARD; PRT; 662 AA.  
 AC P33434;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE 72 KD TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KD GELATINASE)  
 DE (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).  
 GN MMP2.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92218452.  
 RA REPONEN P., SAHLBERG C., HUHTALA P., HURSKAINEN T., THESLEFF I.,  
 RA TRYGGVASON K.;  
 RT "Molecular cloning of murine 72-kDa type IV collagenase and its  
 expression during mouse development.";  
 RL J. BIOL. CHEM. 267:7856-7862(1992).  
 CC -|- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
 IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
 ILE-ALA-GLY-GLN.  
 CC -|- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
 CC -|- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
 CC -|- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
 METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M84324; G198466; -  
 DR PIR; A42496; A42496.

MGD; MGI:97009; MMP2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 DR PFAM; PF00040; In2; 3.  
 DR PFAM; PF00045; hemopexin; 4.  
 DR PFAM; PF00099; zn-protease; 1.  
 DR PFAM; PF00413; matrixin; 1.  
 DR HSP; P08253; IRTG.  
 KW HYDROLASE; METALLOPROTEASE; GLYCOPROTEIN; ZINC; ZMOGEN; CALCIUM;  
 KW COLLAGEN DEGRADATION; EXTRACELLULAR MATRIX; DUPLICATION; SIGNAL.  
 FT SIGNAL 1 29  
 FT PROPEP 30 109  
 FT CHAIN 110 662  
 FT DOMAIN 100 107  
 FT DOMAIN 110 221  
 FT DOMAIN 222 396  
 FT DOMAIN 397 467  
 FT DOMAIN 468 662  
 FT DOMAIN 226 283  
 FT DOMAIN 284 341  
 FT DOMAIN 342 399  
 FT METAL 403 403  
 FT ACT\_SITE 404 404  
 FT METAL 407 407  
 FT METAL 413 413  
 FT CARBOHYD 575 575  
 FT CARBOHYD 644 644  
 FT DISULFID 471 662  
 SQ SEQUENCE 662 AA; 74102 MW; 263E3B37 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 662;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CKFKFK 8  
 DB 291 CKFFPRFQ 298

RESULT 9  
 COG2\_RABIT  
 ID COG2\_RABIT STANDARD; PRT; 662 AA.  
 AC P50757;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE 72 KD TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KD GELATINASE)  
 DE (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).  
 GN MMP2.  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAPANESE WHITE; TISSUE-ARTICULAR JOINT;  
 RX MEDLINE; 96283805.  
 RA MATSUMOTO S., KATO M., WATANABE T., MASUHO Y.;  
 RT "Molecular cloning of rabbit matrix metalloproteinase-2 and its broad  
 expression at several tissues.";  
 RL BIOCHIM. BIOPHYS. ACTA 1307:137-139(1996).  
 CC -|- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
 IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
 ILE-ALA-GLY-GLN.  
 CC -|- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
 CC -|- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
 CC -|- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
 METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; D63579; G944817; -  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 DR PFAM; PF00040; fn2; 3.  
 DR PFAM; PF00045; hemopexin; 4.  
 DR PFAM; PF00099; zn-protease; 1.  
 DR PFAM; PF00413; matrxin; 1.  
 DR HSP; P08253; IRTG.  
 KW HYDROLASE; METALLOPROTEASE; GLYCOPROTEIN; ZINC; ZMOGEN; CALCIUM;  
 KW COLLAGEN DEGRADATION; EXTRACELLULAR MATRIX; DUPLICATION; SIGNAL.  
 FT SIGNAL 1 29  
 FT PROPEP 30 109 ACTIVATION PEPTIDE.  
 FT CHAIN 110 662 72 KD TYPE IV COLLAGENASE.  
 FT DOMAIN 100 107 AUTOINHIBITOR REGION.  
 FT DOMAIN 110 221 COLLAGENASE-LIKE.  
 FT DOMAIN 222 396 COLLAGEN-BINDING.  
 FT DOMAIN 397 467 COLLAGENASE-LIKE.  
 FT DOMAIN 468 662 HEMOPEXIN-LIKE.  
 FT DOMAIN 226 283 FIBRONECTIN TYPE-II.  
 FT DOMAIN 284 341 FIBRONECTIN TYPE-II.  
 FT DOMAIN 342 399 FIBRONECTIN TYPE-II.  
 FT METAL 403 404 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 404 404 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 575 575 POTENTIAL.  
 FT CARBOHYD 644 644 POTENTIAL.  
 FT DISULFID 471 662 BY SIMILARITY.  
 SQ SEQUENCE 662 AA; 73803 MW; 636A4FE4 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 662;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CKFKFKFK 8  
 ||| |. |.  
 Db 291 CKPFRFQ 298

RESULT 10  
 COG2\_RAT STANDARD; PRT; 662 AA.  
 ID COG2\_RAT  
 AC P3436  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE 72 KD TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KD GELATINASE)  
 DE (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).  
 GN MMP2  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP MEDLINE; 93249363.  
 RX SEQUENCE FROM N.A.  
 RA MARTI H.P., MCNEIL L., DAVIES M., MARTIN J., LOVETT D.H.;  
 RT "Homology cloning of rat 72 kDa type IV collagenase: cytokine and  
 RT second-messenger inducibility in glomerular mesangial cells.";  
 RL BIOCHEM. J. 291:441-446(1993).  
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
 CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-I-  
 CC ILE-ALA-GLY-GLN.  
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.

CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X71466; G854415; -  
 DR PIR; S34780; S34780.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 DR PFAM; PF00040; fn2; 3.  
 DR PFAM; PF00045; hemopexin; 4.  
 DR PFAM; PF00099; zn-protease; 1.  
 DR PFAM; PF00413; matrxin; 1.  
 DR HSP; P08253; IRTG.  
 KW HYDROLASE; METALLOPROTEASE; GLYCOPROTEIN; ZINC; ZMOGEN; CALCIUM;  
 KW COLLAGEN DEGRADATION; EXTRACELLULAR MATRIX; DUPLICATION; SIGNAL.  
 FT SIGNAL 1 29  
 FT PROPEP 30 109 ACTIVATION PEPTIDE.  
 FT CHAIN 110 662 72 KD TYPE IV COLLAGENASE.  
 FT DOMAIN 100 107 AUTOINHIBITOR REGION.  
 FT DOMAIN 110 221 COLLAGENASE-LIKE.  
 FT DOMAIN 222 396 COLLAGEN-BINDING.  
 FT DOMAIN 397 467 COLLAGENASE-LIKE.  
 FT DOMAIN 468 662 HEMOPEXIN-LIKE.  
 FT DOMAIN 226 283 FIBRONECTIN TYPE-II.  
 FT DOMAIN 284 341 FIBRONECTIN TYPE-II.  
 FT DOMAIN 342 399 FIBRONECTIN TYPE-II.  
 FT METAL 403 404 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 404 404 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 575 575 POTENTIAL.  
 FT CARBOHYD 644 644 POTENTIAL.  
 FT DISULFID 471 662 BY SIMILARITY.  
 SQ SEQUENCE 662 AA; 74181 MW; B982EC46 CRC32;

Query Match 54.8%; Score 34; DB 1; Length 662;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CKFKFKFK 8  
 ||| |. |.  
 Db 291 CKPFRFQ 298

RESULT 11  
 CPAL\_ECOLI STANDARD; PRT; 61 AA.  
 ID CPAL\_ECOLI  
 AC P03854;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE COPA/INCA PROTEIN (REPA3 PROTEIN).  
 DE COPA OR INCA OR REPA3.  
 GN ESCHERICHIA COLI.  
 OS ESCHERICHIA COLI.  
 OC PLASMID INCFII R100, AND PLASMID INCFII R1.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC ESCHERICHIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC PLASMID-INCFII R100, AND INCFII R1;  
 RX MEDLINE; 81173118.











**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:54:51 ; Search time 31.14 Seconds  
(without alignments)  
12.866 Million cell updates/sec

Title: US-09-124-280A-43  
Perfect score: 62  
Sequence: 1 CKFKFKFKC 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	66.1	596	2 G64987	yejH protein - Esc
2	37	59.7	663	2 S46492	gelatinase A (EC 3
3	37	59.7	190	2 E64522	hypothetical prote
4	37	59.7	191	2 A71985	hypothetical prote
5	37	59.7	1205	2 A55015	bumetanide-sensiti
6	36	58.1	350	2 S34557	hypothetical prote
7	36	58.1	273	2 D71436	hypothetical prote
8	35	56.5	440	2 C70198	conserved hypothet
9	35	56.5	779	2 T01304	hypothetical prote
10	35	56.5	434	2 S61999	hypothetical prote
11	35	56.5	434	2 S55155	probable tyrosine
12	34	54.8	560	1 JC4795	plasma hyaluronan-
13	34	54.8	660	1 A28153	gelatinase A (EC 3
14	34	54.8	61	1 Q0EC71	probable cop prote
15	34	54.8	662	2 A42496	gelatinase A (EC 3
16	34	54.8	662	2 S34780	gelatinase A (EC 3
17	34	54.8	662	2 S70365	gelatinase A (EC 3
18	34	54.8	184	2 S74136	latex proteinase I
19	34	54.8	489	2 C71959	glutamyl tRNA redu
20	34	54.8	489	2 G64349	glutamyl-tRNA redu
21	34	54.8	439	2 T01428	hypothetical prote
22	34	54.8	1938	2 A37361	probable integral
23	34	54.8	650	2 T00094	endostyle-specific
24	34	54.8	170	2 C71199	hypothetical prote
25	33	53.2	364	2 F36470	Wnt-5 protein - mo
26	33	53.2	846	2 A08889	Integrin beta chal
27	33	53.2	430	2 D64151	hypothetical prote
28	33	53.2	322	2 S08451	hypothetical prote
29	33	53.2	201	2 A64737	yadL protein - Esc
30	33	53.2	145	2 S35159	photosystem I chal
31	33	53.2	1135	2 S67794	probable membrane
32	33	53.2	606	2 JW0054	amiloride-sensitiv
33	32.5	52.4	104	2 A39035	ribonuclease-relat
34	32	51.6	1291	1 A48940	bontoxilysin (EC 3
35	32	51.6	2470	2 S57085	1-phosphatidylino
36	32	51.6	209	2 H70189	probable CpGdiacyl
37	32	51.6	1291	2 I40631	non-proteolytic bo
38	32	51.6	336	2 S75947	hypothetical prote
39	32	51.6	485	2 T03131	alkaline exonuclea

ALIGNMENTS

```
RESULT 1
G64987
yejH protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Jun-1998
C:Accession: G64987
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: G64987
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-586 <BLAT>
A:Cross-references: GB:AE000308; GB:U00096; NID:g1788508; PID:g1788511; UWGP:b2184
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yejH
C:Superfamily: unassigned DEAD/H box helicases; DEAD/H box helicase homology
C:Keywords: ATP; P-loop
F:31-343/Domain: DEAD/H box helicase homology <DEAD>
F:31-38/Region: nucleotide-binding motif A (P-loop)
F:113-118/Region: nucleotide-binding motif B
F:117-120/Region: DEXH motif

Query Match 56.1%; Score 41; DB 2; Length 586;
Best Local Similarity 60.0%; Pred. No. 5.7;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CKFKFKFKC 10
| | | | |
Db 425 CDFRFRFKNC 434

RESULT 2
S46492
gelatinase A (EC 3.4.24.24) precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Mar-1998
C:Accession: S46492
R:Aimes, R.T.; French, D.L.; Quigley, J.P.
Biochem. J. 300, 729-736, 1994
A:Title: Cloning of a 72 kDa matrix metalloproteinase (gelatinase) from chicken embryo
A:Reference number: S46492
A:Accession: S46492
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-663 <AIM>
A:Cross-references: EMBL:U07775; NID:g504475; PID:g504476
A:Note: in the authors' translation 205-Asp is shown after residue 201 and, consequen
C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopexin repeat ho
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:67-216,391-443/Domain: matrix metalloproteinase homology #status atypical <MMP>
F:230-271/Domain: fibronectin type II repeat homology <2F1>
F:288-329/Domain: fibronectin type II repeat homology <2F8>
F:346-387/Domain: fibronectin type II repeat homology <2F9>
F:466-663/Domain: hemopexin repeat homology <PXN>
F:99,400,404,410/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:400,404,410/Binding site: zinc, catalytic (His) (active) #status predicted
F:401/Active site: Glu #status predicted
```

RESULT 5  
A55015

RESULT 5  
A55015

hypothetical protein - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 A:Variety: columbia  
 C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 05-Dec-1998  
 C:Accession: D71436  
 P:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
 R.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel  
 avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
 Nature 391, 485-488, 1998  
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
 erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
 C.; Chalwatzis, N.  
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
 A:Reference number: A71400; MUID:98121113  
 A:Accession: D71436  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-273 <BEV>  
 A:Cross-references: GB:297342; NID:g2245031; PID:e327019; PID:g2245042  
 C:Genetics:  
 A:Map position: 4COP9-4G3845

Query Match 58.1%; Score 36; DB 2; Length 273;  
 Best Local Similarity 33.3%; Pred. No. 20;  
 Matches 8; Conservative 1; Mismatches 1; Indels 14; Gaps 1;

Qy 1 CKFKF-----KFKFC 10  
 | | | | |  
 Db 132 CSFKFQPEDLPSKLTNDVFKFC 155  
 | | | | |

RESULT 8  
 C70198  
 conserved hypothetical protein BB0788 - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: C70198  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kervavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: C70198  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-440 <KLE>  
 A:Cross-references: GB:AE001177; GB:AE000783; NID:g2688711; PID:g2688713; TIGR:BB0788  
 A:Experimental source: strain B31

Query Match 56.5%; Score 35; DB 2; Length 440;  
 Best Local Similarity 44.4%; Pred. No. 45;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFKFKFKF 9  
 | : | : | |  
 Db 368 CSYEFYKFK 376  
 | : | : | |

RESULT 9  
 T01304  
 hypothetical protein Tl4P8.6 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
 C:Accession: T01304  
 R:Kalicki, J.; Elliott, G.; Cloud, J.  
 submitted to The EMBL Data Library, May 1998  
 A:Description: The sequence of A. thaliana Tl4P8.  
 A:Reference number: Z14290  
 A:Accession: T01304

A:Status: preliminary; translated from GB/EMBL/DDBU  
 A:Molecule type: DNA  
 A:Residues: 1-779 <KAL>  
 A:Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193291  
 C:Genetics:  
 A:Map position: IV  
 A:Introns: 94/3; 202/3; 254/3; 562/3; 585/3; 610/2; 632/2; 667/1; 727/3  
 A:Note: Tl4P8.6

Query Match 56.5%; Score 35; DB 2; Length 779;  
 Best Local Similarity 66.7%; Pred. No. 74;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KFKFKFKFC 10  
 | | | | |  
 Db 411 KFKFKKAC 419  
 | | | | |

RESULT 10  
 S61999  
 hypothetical protein YPL123c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein LPH4c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 12-Dec-1997  
 C:Accession: S61999  
 R:Schlenstedt, G.; Silver, P.A.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: S61996  
 A:Accession: S61999  
 A:Molecule type: DNA  
 A:Residues: 1-434 <SCH>  
 A:Cross-references: EMBL:U43503; NID:g1163087; PID:g1163091; MIPS:YPL123c  
 C:Genetics:  
 A:Map position: 16L

Query Match 56.5%; Score 35; DB 2; Length 434;  
 Best Local Similarity 75.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KFKFKFKF 9  
 | : | | | |  
 Db 424 KYKFKLKF 431  
 | : | | | |

RESULT 11  
 S55155  
 probable tyrosine phosphatase YNL128w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: probable tyrosine phosphatase N1220; protein JTB434; protein N1872  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 06-Feb-1998  
 C:Accession: S55155; S59260; S63073  
 R:Mallet, L.; Bussereau, F.; Jacquet, M.  
 submitted to the EMBL Data Library, November 1994  
 A:Description: A 43.5 kb fragment of the chromosome XIV.  
 A:Reference number: S55136  
 A:Accession: S55155  
 A:Molecule type: DNA  
 A:Residues: 1-434 <MAL>  
 A:Cross-references: EMBL:Z46843; NID:g861113; PID:g854509  
 R:Mallet, L.; Bussereau, F.; Jacquet, M.  
 Yeast 11, 1195-1209, 1995  
 A:Title: A 43.5 kb segment of yeast chromosome XIV, which contains MFA2, MEP2, CAP/SR  
 A:Reference number: S59241; MUID:96109932  
 A:Accession: S59260  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-434 <MAW>  
 A:Cross-references: EMBL:Z46843; NID:g861113; PID:g854509  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994  
 R:Mallet, L.; Bussereau, F.; Jacquet, M.  
 submitted to the Protein Sequence Database, April 1996

A:Reference number: S63069  
 A:Accession: S63073  
 A:Molecule type: DNA  
 A:Residues: 1-434 <MAF>  
 C:Cross-references: ENBL:271404; NID:gl302063; PID:e239804; PID:gl302064; MIPS:YNL128w  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:TEPI  
 A:Cross-references: SGD:S0005072; MIPS:YNL128w  
 A:Map position: 14L  
 C:Keywords: transmembrane protein  
 F:38-54/Domain: transmembrane #status predicted <TM>

Query Match 56.5%; Score 35; DB 2; Length 434;  
 Best Local Similarity 85.7%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFKFKFK 8  
 Db 268 KFKFKFKR 274

RESULT 12  
 JC4795  
 Plasma hyaluronan-binding protein precursor - human  
 N:Alternate names: hepatocyte growth factor activator-like protein; PHBP  
 N:Contains: serine proteinase (EC 3.4.21.-)  
 C:Species: Homo sapiens (man)  
 C:Date: 15-Oct-1995 #sequence\_revision 16-Aug-1996 #text\_change 13-Nov-1998  
 C:Accession: JC4795  
 R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.  
 J. Biochem. 119, 1157-1165, 1996  
 A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)  
 r activator.  
 A:Reference number: JC4795  
 A:Accession: JC4795  
 A:Molecule type: mRNA  
 A:Residues: 1-360 <CHO>  
 A:Note: parts of this sequence, including the amino ends of the mature chains, were dete

A:Experimental source: plasma  
 C:Genetics:  
 A:Gene: GDB:HABP2; HABP: PHBP; HGFBAL  
 A:Cross-references: GDB:4573962  
 C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; th  
 C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi  
 C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; growth factor; hyaluronic ac  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>  
 F:77-108/Domain: EGF homology <EGF1>  
 F:115-147/Domain: EGF homology <EGF2>  
 F:154-187/Domain: EGF homology <EGF3>  
 F:194-276/Domain: kringle homology <KRI>  
 F:314-550/Domain: trypsin homology <TRY>  
 F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted  
 F:54-207/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F:77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,246  
 F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 54.8%; Score 34; DB 1; Length 560;  
 Best Local Similarity 60.0%; Pred. No. 81;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKFKFKFKC 10  
 Db 178 CPDFKFKFKC 187

RESULT 13  
 A28153  
 gelatinase A (EC 3.4.24.24) precursor - human  
 N:Alternate names: collagenase type IV; matrix metalloproteinase 2 (MMP2); progelatinase

C:Species: Homo sapiens (man)  
 C:Date: 28-Aug-1989 #sequence\_revision 07-Jul-1995 #text\_change 26-Feb-1999  
 C:Accession: A28153; A34202; A42225; A60187; S13858; S29436; A31480; S44432; A61498;  
 R:Collier, I.E.; Wilhelm, S.M.; Eisen, A.Z.; Mamer, B.L.; Grant, G.A.; Seltzer, J.L.  
 J. Biol. Chem. 265, 6579-6587, 1988  
 A:Title: H-ras oncogene-transformed human bronchial epithelial cells (TBE-1) secrete  
 A:Reference number: A28153; MUID:88198218  
 A:Accession: A28153  
 A:Molecule type: mRNA  
 A:Residues: 30-660 <COL>  
 A:Cross-references: GB:J03210; NID:gl180670; PID:gl180671  
 R:Huhtala, P.; Eddy, R.L.; Fan, Y.S.; Byers, M.G.; Shows, T.B.; Tryggvason, K.  
 Genomics 6, 554-559, 1990  
 A:Title: Completion of the primary structure of the human type IV collagenase preproe  
 A:Reference number: A34202; MUID:90228972  
 A:Accession: A34202  
 A:Molecule type: DNA  
 A:Residues: 1-51 <HU2>  
 A:Cross-references: GB:M33789; NID:gl180600; PID:gl180601  
 R:Huhtala, P.; Chow, L.T.; Tryggvason, K.  
 J. Biol. Chem. 265, 11077-11082, 1990  
 A:Title: Structure of the human type IV collagenase gene.  
 A:Reference number: A42225; MUID:90293047  
 A:Accession: A42225  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-51:220-393 <HUH>  
 A:Cross-references: GB:M55593; GB:J05471; NID:gl180614; PID:gl180616  
 A:Note: neither the complete amino acid nor the complete nucleotide sequence is given  
 R:Frisch, S.M.; Reich, R.; Collier, I.E.; Genrich, L.T.; Martin, G.; Goldberg, G.I.  
 Oncogene 5, 75-83, 1990  
 A:Title: Adenovirus E1A represses protease gene expression and inhibits metastasis of  
 A:Reference number: A60187  
 A:Accession: A60187  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-58 <FRI>  
 R:Okada, Y.; Morodomi, T.; Enghild, J.J.; Suzuki, K.; Yasui, A.; Nakanishi, I.; Salve  
 Eur. J. Biochem. 194, 721-730, 1990  
 A:Title: Matrix metalloproteinase 2 from human rheumatoid synovial fibroblasts. Purif  
 A:Reference number: S13858; MUID:91099351  
 A:Accession: S13858  
 A:Molecule type: protein  
 A:Residues: 30-45;110-124 <OKA>  
 R:Crabbe, T.; Ioannou, C.; Docherty, A.J.P.  
 Eur. J. Biochem. 218, 431-438, 1993  
 A:Title: Human progelatinase A can be activated by autolysis at a rate that is concen  
 A:Reference number: S39436; MUID:94094834  
 A:Accession: S39436  
 A:Molecule type: protein  
 A:Residues: 30-44;444-456 <CR2>  
 R:Stetler-Stevenson, W.G.; Kruttsch, H.C.; Wachter, M.P.; Margulies, I.M.K.; Liotta, L  
 J. Biol. Chem. 264, 1353-1356, 1989  
 A:Title: The activation of human type IV collagenase proenzyme. Sequence identificati  
 A:Reference number: A31480; MUID:89109136  
 A:Accession: A31480  
 A:Molecule type: protein  
 A:Residues: 110-123 <STE>  
 R:Crabbe, T.; Smith, B.; O'Connell, J.; Docherty, A.  
 FEBS Lett. 345, 14-16, 1994  
 A:Title: Human progelatinase A can be activated by matrilysin.  
 A:Reference number: S44432; MUID:94252395  
 A:Accession: S44432  
 A:Molecule type: protein  
 A:Residues: 110-115 <CRA>  
 R:Brown, D.; Chwa, M.; Escobar, M.; Kenney, M.C.  
 Exp. Eye Res. 52, 5-16, 1991  
 A:Title: Characterization of the major matrix degrading metalloproteinase of human co  
 A:Reference number: A61498; MUID:91330998  
 A:Accession: A61498  
 A:Molecule type: protein  
 A:Residues: 'X', '31', 'X', '33-46', 'X', '48-50', 'Q' <BRO>  
 A:Experimental source: corneal stroma

R:Ich, Y.; Binner, S.; Nagase, H.  
 Biochem. J. 308, 645-651, 1995  
 A:Title: Steps involved in activation of the complex of pro-matrix metalloproteinase 2 (MMP-2) and tissue inhibitor of metalloproteinases (TIMP-1)  
 A:Reference number: S55327  
 A:Accession: S55327  
 A:Molecule type: protein  
 A:Residues: 110-114 <I>  
 C:Genetics:  
 A:Gene: GDB:MMP2; CLG4; CLG4A  
 A:Cross-references: GDB:120592; OMIM:120360  
 A:Map position: 16q13-16q13  
 A:Introns: 51/3; 127/2; 178/1; 220/1; 278/1; 336/1; 394/1; 446/1; 491/2; 537/1; 590/2; 600/1  
 C:Function:  
 A:Description: proteolytic cleavage of gelatin type I and collagen types IV, V, VII, and XI  
 C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopexin repeat homology  
 C:Keywords: extracellular matrix; fibroblast; glycoprotein; hydrolase; metalloproteinase  
 F:1-29/DNA: signal sequence #status predicted <SIG>  
 F:30-660/Product: progelatinase A #status predicted <PRO>  
 F:30-109/DNA: activation peptide #status predicted <ACT>  
 F:70-219,394-446/DNA: matrix metalloproteinase homology #status atypical <MMP>  
 F:110-660/Product: gelatinase A #status predicted <MAT>  
 F:233-390/Region: collagen binding #status predicted  
 F:233-274/DNA: fibronectin type II repeat homology <2FI>  
 F:291-332/DNA: fibronectin type II repeat homology <2F8>  
 F:349-390/DNA: fibronectin type II repeat homology <2F9>  
 F:463-660/DNA: hemopexin repeat homology <PXN>  
 F:102,403,407,413/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status predicted  
 F:403,407,413/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:404/Active site: Glu #status predicted  
 F:469-660/Disulfide bonds: #status predicted  
 F:573,642/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 54.8%; Score 34; DB 1; Length 660;  
 Best Local Similarity 62.5%; Pred. No. 94;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 CKFKFKFK 8  
 ||| |:  
 Db 291 CKFPFRFQ 298

RESULT 14  
 QOEC71  
 C:Probable cop protein - Escherichia coli plasmids  
 C:Species: Escherichia coli  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 20-Mar-1998  
 C:Accession: B93253; A04483; I64779; A93120; I41108; I54849; I79323  
 R:Rosen, J.; Ryder, T.; Ohtsubo, H.; Ohtsubo, E.  
 Nature 290, 794-797, 1991  
 A:Title: Role of RNA transcripts in replication incompatibility and copy number control  
 A:Reference number: A93253; MUID:81173118  
 A:Accession: B93253  
 A:Molecule type: DNA  
 A:Residues: 1-61 <ROS1>  
 A:Experimental source: plasmid R1  
 A:Accession: A04483  
 A:Molecule type: DNA  
 A:Residues: 1-61 <ROS2>  
 A:Experimental source: plasmid R100  
 R:Ohtsubo, H.; Ryder, T.B.; Maeda, Y.; Armstrong, K.A.; Ohtsubo, E.  
 Adv. Biophys. 21, 115-133, 1986  
 A:Title: DNA replication of the resistance plasmid R100 and its control.  
 A:Reference number: I51821; MUID:86319522  
 A:Accession: I64779  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-61 <RES>  
 A:Cross-references: GB:M26840; NID:g151770; PID:g151772  
 R:Stougaard, P.; Molin, S.; Nordstrom, K.; Hansen, F.G.  
 Mol. Gen. Genet. 181, 116-122, 1981  
 A:Title: The nucleotide sequence of the replication control region of the resistance plasmid R100  
 A:Reference number: A93120; MUID:81172236

A:Accession: A93120  
 A:Molecule type: DNA  
 A:Residues: 1-18, 'W', 20-61 <STO>  
 A:Cross-references: GB:V00326; GB:J01772; NID:g42636; PID:g671862  
 A:Note: this protein is most likely the cop gene product, which controls the copy number of the plasmid  
 R:Womble, D.D.; Sampathkumar, P.; Easton, A.M.; Luckow, V.A.; Rownd, R.H.  
 J. Mol. Biol. 181, 395-410, 1985  
 A:Title: Transcription of the replication control region of the IncFII R-plasmid NRI  
 A:Reference number: I41106; MUID:85160860  
 A:Accession: I41108  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-18, 'W', 20-61 <RES4>  
 A:Cross-references: EMBL:X02302; NID:g42132; PID:g42134  
 R:Brady, G.; Frey, J.; Danbara, H.; Timmis, K.N.  
 J. Bacteriol. 154, 429-436, 1983  
 A:Title: Replication control mutations of plasmid R6-5 and their effects on interaction with the host cell  
 A:Reference number: I54849; MUID:83160806  
 A:Accession: I54849  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-18, 'W', 20-61 <RES>  
 A:Cross-references: EMBL:V00319; NID:g42488; PID:g42489  
 R:Danbara, H.; Brady, G.; Timmis, J.K.; Timmis, K.N.  
 Proc. Natl. Acad. Sci. U.S.A. 78, 4699-4703, 1981  
 A:Title: Regulation of DNA replication: 'target' determinant of the replication control region  
 A:Reference number: I58974; MUID:82060121  
 A:Accession: I79323  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-18, 'W', 20-61 <RES3>  
 A:Cross-references: EMBL:V00318; NID:g42485; PID:g42487  
 C:Genetics:  
 A:Genome: plasmid  
 C:Complex: homodimer  
 C:Function:  
 A:Description: controls the copy number in gene replication  
 C:Superfamily: cop protein  
 C:Keywords: plasmid copy control

Query Match 54.8%; Score 34; DB 1; Length 61;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Oy 1 CKFKFKFKFC 10  
 | ||| |  
 Db 45 CPERFKTSFC 54

RESULT 15  
 A42496  
 gelatinase A (EC 3.4.24.24) precursor - mouse  
 N:Alternate names: collagenase type IV, 72K  
 C:Species: Mus musculus (house mouse)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Mar-1998  
 C:Accession: A42496  
 R:Reponen, P.; Sahlberg, C.; Huhtala, P.; Hurskainen, T.; Thesleff, I.; Tryggvason, K.  
 J. Biol. Chem. 267, 7856-7862, 1992  
 A:Title: Molecular cloning of murine 72-kDa type IV collagenase and its expression du  
 A:Reference number: A42496; MUID:92218452  
 A:Accession: A42496  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-662 <REP>  
 A:Cross-references: GB:M84324; NID:g198465; PID:g198466  
 A:Note: sequence extracted from NCBI backbone (NCBI:96943, NCBIP:96945)  
 C:Superfamily: gelatinase A; fibronectin type II repeat homology; hemopexin repeat ho  
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:233-274/DNA: fibronectin type II repeat homology <2FI>  
 F:291-332/DNA: fibronectin type II repeat homology <2F8>  
 F:349-390/DNA: fibronectin type II repeat homology <2F9>  
 F:465-662/DNA: hemopexin repeat homology <PXN>

F:102,403,407,413/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
F:403,407,413/Binding site: zinc, catalytic (His) (active) #status predicted  
F:404/Active site: Glu #status predicted

Query Match 54.8%; Score 34; DB 2; Length 662;  
Best Local Similarity 62.5%; Pred. No. 95;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKFKFK 8  
||| |:  
DB 291 CKFPFRFQ 298

Search completed: September 7, 1999, 23:54:52  
Job time: 1903 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:23:07 ; Search time 25.2 Seconds  
(without alignments)  
3.916 Million cell updates/sec

Title: US-09-124-280A-43  
Perfect score: 62  
Sequence: 1 CKFKFKFKC 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	2	US-08-456-112B-43
2	44	71.0	10	1	US-08-097-830E-5
3	44	71.0	10	2	US-08-456-112B-5
4	36	58.1	10	1	US-08-097-830E-15
5	36	58.1	10	2	US-08-456-112B-15
6	36	58.1	314	2	US-08-460-309-19
7	36	58.1	314	2	US-08-125-077-19
8	33	53.2	10	1	US-08-311-611A-165
9	33	53.2	10	1	US-08-372-783-165
10	33	53.2	10	1	US-08-372-105-165
11	33	53.2	10	1	US-08-306-473A-165
12	33	53.2	10	2	US-08-621-803-124
13	33	53.2	10	2	US-08-485-445A-165
14	33	53.2	10	3	PCT-US95-00498-165
15	33	53.2	10	3	PCT-US95-00556-165
16	32.5	52.4	104	1	US-08-283-971-1
17	32.5	52.4	104	1	US-07-921-619-1
18	32.5	52.4	104	1	US-08-467-955-1
19	32.5	52.4	104	1	US-08-467-955-2
20	32	51.6	133	1	US-07-893-929A-9
21	32	51.6	351	1	US-08-415-751-41
22	32	51.6	133	3	PCT-US92-10344-9
23	31	50.0	48	1	US-07-965-674-7
24	31	50.0	482	1	US-08-135-510-5
25	31	50.0	504	1	US-08-135-511-28
26	31	50.0	482	1	US-08-483-852-5
27	31	50.0	504	1	US-08-483-852-8
28	31	50.0	504	1	US-08-361-458-3
29	31	50.0	221	1	US-07-949-812-2
30	31	50.0	192	1	US-07-949-812-3
31	31	50.0	482	1	US-08-477-953-5
32	31	50.0	504	1	US-08-477-953-8
33	31	50.0	493	1	US-08-362-512A-4
34	31	50.0	48	1	US-08-462-661A-14
35	31	50.0	48	1	US-08-462-661A-16
36	31	50.0	504	2	US-08-187-453-28
37	31	50.0	359	2	US-08-748-485-5
38	31	50.0	482	2	US-08-477-952-5
39	31	50.0	504	2	US-08-477-952-8

40 31 50.0 48 3 PCT-US93-09523-7 Sequence 7, Appl  
41 30.5 49.2 10 1 US-08-097-830E-18 Sequence 18, Appl  
42 30.5 49.2 10 2 US-08-456-112B-18 Sequence 18, Appl  
43 30 48.4 213 1 US-08-104-445-3 Sequence 3, Appl  
44 30 48.4 296 1 US-08-261-662-2 Sequence 2, Appl  
45 30 48.4 553 1 US-08-328-322-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-08-456-112B-43  
; Sequence 43, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: PORIO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-43

Query Match 100.0%; Score 62; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKFKFKFKC 10  
| | | | | | | | | |  
Db 1 CKFKFKFKC 10

RESULT 2  
US-08-097-830E-5  
; Sequence 5, Application US/08097830E  
; Patent No. 5652211  
; GENERAL INFORMATION:  
; APPLICANT: PORIO, Massimo  
; TITLE OF INVENTION: Peptides For Neutralizing The  
; TITLE OF INVENTION: Toxicity of Lipid A  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan, P.C.

STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8998  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-097-830E-5

Query Match 71.0%; Score 44; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFKFKFKF 9  
Db 1 KFKFKFKF 8

RESULT 3  
US-08-456-112B-5  
Sequence 5, Application US/08456112B  
Patent No. 5834430  
GENERAL INFORMATION:  
APPLICANT: Porrio, Massimo  
TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: LEADING EDGE 486  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,112B  
FILING DATE: May 31, 1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-004  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 302-8989  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
US-08-456-112B-5

Query Match 71.0%; Score 44; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFKFKFKF 9  
Db 1 KFKFKFKF 8

RESULT 4  
US-08-097-830E-15  
Sequence 15, Application US/08097830E  
Patent No. 5652211  
GENERAL INFORMATION:  
APPLICANT: Porrio, Massimo  
TITLE OF INVENTION: Peptides For Neutralizing The  
TITLE OF INVENTION: Toxicity of Lipid A  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: DOS  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,830E  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Costigan, James V.  
REGISTRATION NUMBER: 25,669  
REFERENCE/DOCKET NUMBER: 576-003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 302-8998  
TELEFAX: (212) 302-8998  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: circular  
FEATURE:  
OTHER INFORMATION: sulfide bond between Cys and  
OTHER INFORMATION: Cys  
US-08-097-830E-15

Query Match 58.1%; Score 36; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKFKF 6  
Db 4 CKFKFKF 9

RESULT 5  
US-08-456-112B-15  
; Sequence 15, Application US/08456112B  
; Patent No. 5834430  
; GENERAL INFORMATION:  
; APPLICANT: POIRIO, Massimo  
; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hedman, Gibson & Costigan  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: LEADING EDGE 486  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.112B  
; FILING DATE: May 31, 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Costigan, James V.  
; REGISTRATION NUMBER: 25,669  
; REFERENCE/DOCKET NUMBER: 576-004  
; TELEPHONE: (212) 302-8989  
; TELEFAX: (212) 302-8998  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-08-456-112B-15

Query Match 58.1%; Score 36; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKFK 6  
| | | | |  
Db 4 CKFKFK 9

RESULT 6  
US-08-460-309-19  
; Sequence 19, Application US/08460309  
; Patent No. 5837496  
; GENERAL INFORMATION:  
; APPLICANT: Engvall, Eva  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; TITLE OF INVENTION: Fragments and Uses Thereof  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,309  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/125,077  
; FILING DATE: 22-SEP-1993  
; APPLICATION NUMBER: US PCT/US 94/10730  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/472,319  
; FILING DATE: 30-JAN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/919,951  
; FILING DATE: 27-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9721  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-460-309-19

Query Match 58.1%; Score 36; DB 2; Length 314;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CKFKFKFKC 10  
| | | | | | | | | |  
Db 198 CKINFAYC 207

RESULT 7  
US-08-125-077-19  
; Sequence 19, Application US/08125077  
; Patent No. 5872231  
; Patent No. 5872231 5840863  
; GENERAL INFORMATION:  
; APPLICANT: Engvall, Eva  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; TITLE OF INVENTION: Fragments and Uses Thereof  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/125,077  
; FILING DATE: 22-SEP-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US 94/10730  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/472,319  
; FILING DATE: 30-JAN-1990

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/919,951  
; FILING DATE: 27-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9721  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-125-077-19

Query Match 58.1%; Score 36; DB 2; Length 314;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CKFKFKFKFC 10  
Db 198 CKINFAFYC 207

## RESULT 8

US-08-311-611A-165  
; Sequence 165, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 165:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid

; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "BPI.169"  
US-08-311-611A-165  
  
Query Match 53.2%; Score 33; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CKFKFKFKFC 10  
Db 1 CKWKAQRFC 10  
  
RESULT 9  
US-08-372-783-165  
; Sequence 165, Application US/08372783  
; Patent No. 5578572  
; GENERAL INFORMATION:  
; APPLICANT: Horwitz, Arnold H.  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and  
; TITLE OF INVENTION: Materials  
; NUMBER OF SEQUENCES: 237  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372,783  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 165:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "XMP.169"  
US-08-372-783-165

Query Match

53.2%; Score 33; DB 1; Length 10;

Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFKFKFC 10  
||| : ||  
Db 1 CKWAKKFC 10

## RESULT 10

US-08-372-105-165  
; Sequence 165, Application US/08372105  
; Patent No. 5627153  
; GENERAL INFORMATION:  
; APPLICANT: Little, Roger G.  
; APPLICANT: Lim, Edward  
; APPLICANT: Lambert, Lewis H.  
; APPLICANT: Scannon, Patrick J.  
; TITLE OF INVENTION: Anti-Fungal Materials and Methods  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/372.105  
; FILING DATE:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273.540  
; FILING DATE: 11-JUL-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209.762  
; FILING DATE: 11-MAR-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183.222  
; FILING DATE: 14-JAN-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 165:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "XMP.169"

US-08-372-105-165

Query Match 53.2%; Score 33; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFKFKFC 10  
||| : ||  
Db 1 CKWAKKFC 10

## RESULT 11

Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

US-08-306-473A-165  
; Sequence 165, Application US/08306473A  
; Patent No. 5652332  
; GENERAL INFORMATION:

; APPLICANT: Little, Roger G.  
; TITLE OF INVENTION: Biologically Active Peptides from  
; TITLE OF INVENTION: Functional Domains of Bactericidal/  
; TITLE OF INVENTION: Permeability-Increasing Protein and  
; TITLE OF INVENTION: Uses Thereof  
; NUMBER OF SEQUENCES: 226  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: Suite 3000, 10 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306.473A  
; FILING DATE:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209.762  
; FILING DATE: 11-MAR-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183.222  
; FILING DATE: 14-JAN-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: McDonnell, John J.  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 93,1133-  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; INFORMATION FOR SEQ ID NO: 165:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "BPI.169"

US-08-306-473A-165

Query Match 53.2%; Score 33; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFKFKFC 10  
||| : ||  
Db 1 CKWAKKFC 10

## RESULT 12

US-08-621-803-124  
; Sequence 124, Application US/08621803  
; Patent No. 5851802  
; GENERAL INFORMATION:

; APPLICANT: Better, Marc D.  
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of  
; TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois

US-08-621-803-124

```

; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621.803
; FILING DATE: 22-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.169 reduced"
; NAME/KEY: Modified-site
; LOCATION: C-Terminus
; OTHER INFORMATION: /label= Amidation
; OTHER INFORMATION: /note= "The C-terminus is Amidated."
US-08-621-803-124

```

```

Query Match 53.2%; Score 33; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.3;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 CKFKFKFKFC 10
||| : |||
Db 1 CKWKAQKRF 10

```

```

RESULT 13
US-08-485-445A-165
; Sequence 165, Application US/08485445A
; Patent No. 5856438
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; TITLE OF INVENTION: Biologically Active Peptides from
; TITLE OF INVENTION: Functional Domains of Bactericidal/
; TITLE OF INVENTION: Permeability-Increasing Protein and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 226
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: Suite 3400, 500 West Madison Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.445A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 110180S08/100-224.P4.C1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-707-8889
; TELEFAX: 312-707-9155
; INFORMATION FOR SEQ ID NO: 165:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: circular
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.169"
US-08-485-445A-165

Query Match 53.2%; Score 33; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 5.3;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CKFKFKFKFC 10
||| : |||
Db 1 CKWKAQKRF 10

RESULT 14
PCT-US95-00498-165
; Sequence 165, Application PC/TUS9500498
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and
; TITLE OF INVENTION: Materials
; NUMBER OF SEQUENCES: 237
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/00498
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27129/32415
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 165:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids

```

;  
; TYPE: amino acid  
; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "XMP.169"  
PCT-US95-00498-165

Query Match 53.2%; Score 33; DB 3; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFKFKFKC 10  
Db 1 CKWKAQKFC 10

RESULT 15  
PCT-US95-00656-165  
; Sequence 165, Application PC/TUS9500656  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Anti-Fungal Materials and Methods  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/00656  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,540  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/209,762  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,222  
; FILING DATE: 14-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rin-Laures, Li-Hsien  
; REGISTRATION NUMBER: 33,547  
; REFERENCE/DOCKET NUMBER: 27129/32415  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 165:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: "XMP.169"  
PCT-US95-00656-165

Query Match 53.2%; Score 33; DB 3; Length 10;  
Best Local Similarity 50.0%; Pred. No. 5.3;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CKFKFKFKC 10

Db 1 CKWKAQKFC 10

Search completed: September 7, 1999, 23:23:08  
Job time: 912 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:18:21 ; Search time 40.95 Seconds  
(without alignments)  
5.784 Million cell updates/sec

Title: US-09-124-280A-43  
Perfect score: 62  
Sequence: 1 CKFKFKFKC 10  
Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	10	1 W21631	Antibiotic potenti
2	44	71.0	10	1 R71776	Peptide neutralisi
3	44	71.0	10	1 W21593	Antibiotic potenti
4	37	59.7	663	1 W41111	Chicken matrix met
5	37	59.7	194	1 W98780	H. pylori GHPO 119
6	36	58.1	10	1 R71786	Peptide neutralisi
7	36	58.1	10	1 W21603	Antibiotic potenti
8	35	56.5	18	1 W52492	Cyclic peptide of
9	34	54.8	631	1 P91139	Human type IV coll
10	34	54.8	660	1 R06420	Type IV collagenas
11	34	54.8	631	1 R07969	Complete type IV c
12	34	54.8	631	1 P36143	Sequence of human
13	34	54.8	429	1 W41112	Human matrix metal
14	34	54.8	631	1 W41226	Human mature matr
15	34	54.8	449	1 W98821	H. pylori GHPO 142
16	33	53.2	10	1 R81243	Anti-fungal BPI pe
17	33	53.2	10	1 R87817	BPI-169 for use in
18	33	53.2	10	1 R90153	Cys-bounded BPI-(9
19	33	53.2	10	1 R82324	BPI-169, domain I
20	33	53.2	10	1 R78077	BPI protein segmen
21	33	53.2	10	1 W06012	Recombinant BPI pe
22	33	53.2	10	1 W43680	Bactericidal/perme
23	33	53.2	10	1 W63463	Human BPI protein
24	32.5	52.4	104	1 R12344	Protein with activ
25	32.5	52.4	104	1 R73033	ONCONASE (pharmace
26	32.5	52.4	104	1 W00736	Protein derived fr
27	32.5	52.4	104	1 W14065	Onconase (RTM) pro
28	32.5	52.4	9	1 W21628	Antibiotic potenti
29	32.5	52.4	104	1 W06544	Antitumour protein
30	32.5	52.4	104	1 W18224	Antitumour generic
31	32.5	52.4	104	1 W06543	Antitumour protein
32	32.5	52.4	379	1 W35126	R. pipiens recombi
33	32.5	52.4	355	1 W35129	R. pipiens recombi
34	32.5	52.4	358	1 W35130	R. pipiens recombi
35	32.5	52.4	366	1 W35132	R. pipiens recombi
36	32.5	52.4	355	1 W35133	R. pipiens recombi
37	32.5	52.4	254	1 W35135	R. pipiens recombi
38	32.5	52.4	251	1 W35134	R. pipiens recombi
39	32.5	52.4	111	1 W35121	R. pipiens recombi
40	32.5	52.4	106	1 W35122	R. pipiens recombi
41	32.5	52.4	105	1 W35123	R. pipiens recombi
42	32.5	52.4	83	1 W35119	R. pipiens clone R
43	32.5	52.4	355	1 W35125	R. pipiens recombi

ALIGNMENTS

RESULT 1  
W21631  
ID W21631 standard; peptide; 10 AA.  
AC W21631:  
DT 26-AUG-1997 (first entry)  
DE Antibiotic potentiating peptide #43.  
KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
KW permeability; outer bacterial membrane.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..10  
PN W09638163-Al.  
PD 05-DEC-1996.  
PF 29-MAY-1996; E02313.  
PR 31-MAY-1995; US-456112.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M, Varra M;  
DR WPI; 97-034095/03.  
PT Potentiating activity of antibiotic with peptide contg. cationic  
PT amino acid sequence - reduces dose of antibiotic required  
PS Claim 47: Page 29: 37pp: English.  
CC The sequences given in W21589-633 represent peptides which act to  
CC potentiate the activity of an antibiotic when they are co-administered  
CC with the antibiotic. Compositions containing these peptides are used  
CC to treat or prevent microbial infections. These peptides bind to  
CC lipopolysaccharide on the bacteria so may increase permeability of  
CC the outer bacterial membrane to the antibiotic, allowing a reduction  
CC in the dose of antibiotic required by 10-90% of the normal dose for  
CC in vivo or in vitro application. Any toxic side effects are  
CC correspondingly reduced.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 62; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0001;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CKFKFKFKFC 10  
Db 1 CKFKFKFKFC 10

RESULT 2  
R71776  
ID R71776 standard; peptide; 10 AA.  
AC R71776:  
DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
OS Synthetic.  
PN W09503327-A.  
FD 02-FEB-1995.  
PF 21-JUL-1994; E02413.  
PR 26-JUL-1993; US-097830.  
PA (BIOS-) BIOSYNTH SRL.  
PI Porro M;  
DR WPI; 95-075190/10.  
PT New peptide(s) for neutralising LPS endotoxin - comprising  
PT repeating units of a basic aminoacid or basic and hydrophobic  
PT amino acids  
PS Claim 7: Page 20: 26pp: English.  
CC New peptides are claimed which are linear or cyclic peptides of formula:  
CC (A)n, where A is the cationic amino acid Lys or Arg and n is 7-10;  
CC (AB)m, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu,  
CC Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)p, where A is  
CC Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or

CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic  
 CC activity. Hence they can be used therapeutically to treat septic shock  
 CC and also in vitro to detoxify vaccines, drug solutions, injectable  
 CC nutrient solutions, etc.  
 CC The present sequence is a specifically claimed example of the new  
 CC peptides.  
 SQ Sequence 10 AA;

Query Match 71.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.079;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFKFKFKF 9  
 |||||  
 Db 1 KFKFKFKF 8

## RESULT 3

ID W21593 standard; peptide: 10 AA.  
 AC W21593;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #5.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN W09638163-Al.  
 PD 05-DEC-1996.  
 PR 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 PS Claim 9; Page 24; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 10 AA;

Query Match 71.0%; Score 44; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.079;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KFKFKFKF 9  
 |||||  
 Db 1 KFKFKFKF 8

## RESULT 4

ID W4111 standard; Protein: 663 AA.  
 AC W4111;  
 DT 08-JUN-1998 (first entry)  
 DE Chicken matrix metalloproteinase-2.  
 KW Matrix metalloproteinase-2; MMP-2; chMMP-2; chicken;  
 KW angiogenesis; inhibitor; antagonist; integrin alpha-v beta-3;  
 KW vitronectin receptor; rheumatoid arthritis; tumour; metastasis;  
 KW diabetic retinopathy; macular degeneration; restenosis; therapy.  
 OS Gallus sp.  
 FH Key Location/Qualifiers  
 FT Peptide 1..26  
 FT /label- Sig\_peptide

PN W09745137-Al.  
 PD 04-DEC-1997.  
 PR 30-MAY-1997; U09158.  
 PR 31-MAY-1996; US-018733.  
 PR 31-MAY-1996; US-015869.  
 PA (SCRI) SCRIPPS RES INST.  
 PI Brooks P, Cheres DA;  
 DR WPI: 98-032334/03.  
 DR N-PSDB; W03995.  
 PT Packaging material containing polypeptide antagonist of alphav,  
 PT beta3 integrin - used for inhibition of angiogenesis, and for  
 PT treating tumours, inflammation, eye diseases etc.  
 PS Disclosure; Page 163-167; 234pp; English.  
 CC This protein sequence comprises chicken matrix metalloproteinase-2  
 CC (chMMP-2). The invention relates to the discovery that angiogenesis  
 CC is mediated by the specific vitronectin receptor alpha-v beta-3,  
 CC and that inhibition of alpha-v beta-3 function inhibits  
 CC angiogenesis. Claimed antagonists of alpha-v beta-3 include  
 CC C-terminal fragments (see W41083-94) of human or chicken MMP-2. An  
 CC MMP-2 fragment can be obtained by recombinant DNA methods, such as  
 CC PCR amplification of the chMMP-2 coding region, cloning into e.g.  
 CC pGEX-3X, and expression in E. coli as a fusion protein with  
 CC glutathione-S-transferases. The antagonists can be used to inhibit  
 CC angiogenesis in inflamed tissue (for treatment of arthritis or  
 CC rheumatoid arthritis), in solid tumours or metastases (particularly  
 CC to induce regression or inhibit tumour growth), and in ocular  
 CC disorders such as diabetic retinopathy and macular degeneration, as  
 CC well as to treat restenosis (all claimed).  
 SQ Sequence 663 AA;

Query Match 59.7%; Score 37; DB 1; Length 663;  
 Best Local Similarity 75.0%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKFKFK 8  
 |||||  
 Db 288 CKFKFKFK 295

## RESULT 5

ID W98780 standard; Protein: 194 AA.  
 AC W98780;  
 DT 31-MAR-1999 (first entry)  
 DE H. pylori GHPO 1193 protein.  
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 OS Helicobacter pylori.  
 PN W09843478-Al.  
 PD 08-OCT-1998.  
 PR 01-APR-1998; U06371.  
 PR 29-JUL-1997; US-902615.  
 PR 01-APR-1997; US-833457.  
 PR 24-JUN-1997; US-881227.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 DR WPI: 98-542293/46.  
 DR N-PSDB; X14499.  
 PT New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 PS Claim 8; Page 1691-1692; 2054pp; English.  
 CC This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 SQ Sequence 194 AA;

Query Match 59.7%; Score 37; DB 1; Length 194;  
 Best Local Similarity 77.8%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CKFKFKFC 10  
 II II III  
 Db 3 KFLFKQKFC 11

## RESULT 6

R71786  
 ID R71786 standard; peptide; 10 AA.

AC R71786;  
 DT 01-OCT-1995 (first entry)  
 DE Peptide neutralising toxicity of Lipid A.  
 KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.  
 OS Synthetic.

FH Key Location/Qualifiers

FT disulfide\_bond 4..10

PN WO9503327-A.

PD 02-FEB-1995.

PF 21-JUL-1994; E02413.

PR 26-JUL-1993; US-097830.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M;

DR WPI; 95-075190/10.

PT New peptide(s) for neutralising LPS endotoxin - comprising repeating units of a basic aminoacid or basic and hydrophobic amino acids

PS Claim 17; Page 21; 26pp; English.

CC New peptides are claimed which are linear or cyclic peptides which include units of formula: (A)<sub>n</sub>, where A is the cationic amino acid Lys or Arg and n is 7-10; (AB)<sub>m</sub>, where A is Lys or Arg, B is the hydrophobic amino acid Val, Leu, Ile, Tyr, Phe or Trp, and m is 3 or greater; and (ABC)<sub>p</sub>, where A is Lys or Arg, B and C are Val, Leu, Ile, Tyr, Phe or Trp, and p is 2 or greater.

CC The peptides bind to Lipid A of endotoxins and do not exhibit haemolytic activity. Hence they can be used therapeutically to treat septic shock and also in vitro to detoxify vaccines, drug solutions, injectable nutrient solutions, etc.

CC The present sequence is a specifically claimed example of the new peptides.

CC Sequence 10 AA;

Query Match 58.1%; Score 36; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKFK 6  
 I I I I I  
 Db 4 CKFKFK 9

## RESULT 7

W21603  
 ID W21603 standard; peptide; 10 AA.

AC W21603;

DT 26-AUG-1997 (first entry)

DE Antibiotic potentiating peptide #15.

KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;

KW permeability; outer bacterial membrane.

OS Synthetic.

FH Key Location/Qualifiers

FT disulfide\_bond 4..10

PN WO9638163-A1.

PD 05-DEC-1996.

PF 29-MAY-1996; E02313.

PR 31-MAY-1995; US-456112.

PA (BIOS-) BIOSYNTH SRL.

PI Porro M, Varra M;

DR WPI; 97-034095/03.

PT Potentiating activity of antibiotic with peptide contg. cationic amino acid sequence - reduces dose of antibiotic required  
 PS Claim 19; Page 25; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to potentiate the activity of an antibiotic when they are co-administered with the antibiotic. Compositions containing these peptides are used to treat or prevent microbial infections. These peptides bind to lipopolysaccharide on the bacteria so may increase permeability of the outer bacterial membrane to the antibiotic, allowing a reduction in the dose of antibiotic required by 10-90% of the normal dose for in vivo or in vitro application. Any toxic side effects are correspondingly reduced.  
 CC Sequence 10 AA;  
 SQ

Query Match 58.1%; Score 36; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CKFKFK 6  
 I I I I I  
 Db 4 CKFKFK 9

## RESULT 8

W52492

ID W52492 standard; peptide; 18 AA.

AC W52492;

DT 01-JUL-1998 (first entry)

DE Cyclic peptide of the invention.

KW Loop region; cyclic peptide; antimicrobial; disinfectant; therapy;

KW preservative; amphipathic anti-parallel beta-sheet region; plant disease. Synthetic.

FH Key Location/Qualifiers

FT Misc\_difference 7 /note= "D-form residue"

PN WO9803192-A1.

PD 29-JAN-1998.

PF 23-JUL-1997; U12974.

PR 24-JUL-1996; US-685589.

PA (INTR-) INTRABIOTICS PHARM INC.

PI Chang C, Chen J, Gu L;

DR WPI; 98-120472/11.

PT New cyclic peptide(s) with antimicrobial activity - contain amphipathic beta-sheet, loop and beta-turn regions, have better

PT activity, bio-availability and protease resistance than linear

PT analogues

PS Claim 9; Page 154; 160pp; English.

CC This sequence is an example of a cyclic peptides (I) of the invention, which has: (a) an amphipathic anti-parallel beta-sheet region (SR), a loop region (LR) and a beta-turn region (TR); (b) a net positive charge at physiological pH; and (c) at least one basic amino acid (aa) in LR or TR. (I) are broad spectrum antimicrobials, specifically for use against E. coli, Pseudomonas aeruginosa, methicillin-resistant Staphylococcus aureus (MRSA), vancomycin-resistant Enterococcus faecium and penicillin-resistant Streptococcus pneumoniae. More generally they are active against Gram-positive or -negative bacteria, fungi, yeast and protozoa. Apart from clinical uses, (I) are also used as disinfectants and preservatives for medical equipment, foods, cosmetics etc., also for treatment of plant diseases. Compared with non-cyclised analogues (i.e. tachyplesin and protegrin type peptides), (I) and are more effective, with better bioavailability and/or serum half-life (increased resistance to proteolysis). They are more suitable for oral administration, can be used at lower doses and are unlikely to induce development of resistant strains.

CC Sequence 18 AA;

Query Match 56.5%; Score 35; DB 1; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 3.8;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CKFKFKFKFC 10

Db 1:1 1:1  
5 CKFPFCVRF 14

## RESULT 9

ID P91139 standard; protein: 631 AA.

AC P91139;  
DE Human type IV collagenase (gelatinase).  
KW Human type IV collagenase; gelatinase; hypertrophic scars; keloids;  
KW intervertebral disc disease; extracellular matrix metalloproteinase;  
KW bronchial epithelial cells; TBE-1 cells; pGell86.2; type II motif;  
KW fibronectin; collagen-binding domain.  
OS Homo sapiens

FT Key Location/Qualifiers

FT domain 1..192

FT domain 193..367

FT duplication 368..631

FT duplication 197..254

FT duplication 255..312

FT duplication 313..368

PN GB2209526-A.

PD 17-MAY-1989.

PF 02-SEP-1988; 020803.

PR 04-SEP-1987; US-093421.

PA (UNIW) Washington University.

PI Eisen ZA, Goldberg GI;

DR WPI; 89-147011/20.

PT DNA encoding human type IV collagenase (gelatinase) - for use in the treatment of hypertrophic scars, keloids and intervertebral disc disease.

PS Claim 2; Fig6; 36pp; English.

CC Human type IV collagenase (gelatinase). Protein source was H-ras transformed human bronchial epithelial cells (TBE-1). The sequence was determined from clone pGel 186.2 which represents almost the full mRNA sequence. Feature 1 is the N-terminal domain, 1; feature 2 is a middle domain, 11, which is organised into 3 x 58 amino acid long head to tail repeats (features 4,5, and 6). These show homology to the type II motif collagen binding domain of fibronectin. Feature 3 is the C-terminal domain. The enzyme could be used in the treatment of hypertrophic scars, keloids, and intervertebral disc disease.

CC See also N91700.

SQ Sequence 631 AA;

Query Match 54.8%; Score 34; DB 1; Length 631;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFKPKFK 8

111 1:1

Db 262 CKFPFRFQ 269

## RESULT 10

ID R06420 standard; protein: 660 AA.

AC R06420;

DE 13-DEC-1990 (first entry)

DV Type IV collagenase cDNA product.

KW hypertrophic scars; keloids; intervertebral disc disease; ds.

OS Homo sapiens.

PN US4923818-A.

PD 08-MAY-1990.

PF 15-MAY-1989; 352069.

PR 15-MAY-1989; US-352069.

PA (UNIW) UNIV OF WASHINGTON.

PI Goldberg GL, Eisen AZ;

DR WPI; 90-245482/32.

DR N-PSDB; Q05620.

PT Recombinant human type IV collagenase - used in treatment of

PT hypertrophic scars, keloids and intervertebral disc disease

PS Claim 3; Fig 9; 23pp; English.

CC cDNA clone enables production of type IV collagenase, useful in  
CC catalysing cleavage of extracellular matrix macromolecules, and  
CC in treatment of hypertrophic scars, keloids and intervertebral disc  
CC disease.  
SQ Sequence 660 AA;

Query Match 54.8%; Score 34; DB 1; Length 660;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CKFKPKFK 8

111 1:1

Db 291 CKFPFRFQ 298

## RESULT 11

R07969

ID R07969 standard; protein: 631 AA.

AC R07969;

DE 16-JAN-1991 (first entry)

DV Complete type IV collagenase.

KW Type IV collagenase; peptide fragments; metalloproteinase detection;

KW antibodies; metalloproteinase inhibition; angiogenesis; arthritis;

KW tumour growth; metastasis; granulomatous inflammatory conditions;

KW sarcoidosis.

OS Homo sapiens.

PH Key

FT peptide 1..18

FT /label=1

FT peptide 19..33

FT /label=2

FT peptide 26..42

FT /label=3

FT protein 34..50

FT /label=4

FT peptide 51..66

FT /label=5

FT peptide 67..80

FT /label=6

FT peptide 67..89

FT /label=7

FT peptide 69..75

FT /label=8

FT peptide 75..94

FT /label=9

FT peptide 141..150

FT /label=10

FT peptide 299..307

FT /label=11

FT peptide 308..318

FT /label=12

FT peptide 344..368

FT /label=13

FT peptide 371..386

FT /label=14

FT peptide 372..375

FT /label=15

FT peptide 472..491

FT /label=16

US7317407-A.

21-AUG-1990.

PD 01-MAR-1989; 317407.

PR 01-MAR-1989; US-317407.

PR 26-FEB-1990; US-488460.

PA (USSH) NAT CANCER INST.

PI Liotta LA, Stetler-Stevenson W, Krutzsch H;

DR WPI; 90-290093/38.

PT New type-IV collagenase peptide fragments - used for

PT metallo-proteinase detection and inhibition and for producing

PT antibodies for enzyme detection

PS Disclosure; Fig 1; 42pp; English.

CC Type IV procollagenase was purified from human A2058 melanoma cells.

CC The complete amino acid sequence was determined (see also Hoyhtya, M. et al, (1988) FEBS letters 233, 109-113). Based on this sequence, CC peptides were synthesised (see features) having homology with a CC histidine contg. domain at residues 371-386, a cysteine contg. domain CC at residues 200-370, the 80 residue amino terminus or a CC region 159 residues from the carboxy terminus. These regions CC correspond to the domain of the enzyme involved in enzyme CC activation and interaction of the enzyme with the substrate. CC The peptides are useful in metalloproteinase detection and CC inhibition. They can be used in the treatment of inappropriate CC angiogenesis, arthritis, tumour growth, invasion and metastasis and CC granulomatous inflammatory conditions such as sarcoidosis. CC The peptides can be used to produce antibodies. CC Peptide 6, at concn. of 0.1 mM inhibited 80% of the enzyme CC activity. CC See also US7494796-A and WO9010228. CC Sequence 631 AA;

Query Match 54.8%; Score 34; DB 1; Length 631;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKFKFK 8  
||| |::: 269  
Db 262 CKFFFRFQ 269

## RESULT 12

ID P96143 standard; Protein; 631 AA.  
AC P96143;  
DT 09-MAY-1991 (first entry)  
DE Sequence of human type IV collagenase (gelatinase) in  
DE pGEL 186.2.  
KW Hypertrophic scar; keloid; intervertebral disc disease; enzyme.  
CC Homo sapiens.  
OS Homo sapiens.  
PN G82209528-A.  
PD 17-MAY-1989.  
PF 02-SEP-1988; 20803.  
PR 04-SEP-1987; US-093421.  
PA (UNIM) UNIV OF WASHINGTON.  
PI Elsen AZ, Goldberg GI;  
DR WPI; 89-147011/20.  
DR N-PSDB; N91700.  
PT DNA encoding human type IV collagenase (gelatinase) - for use in the  
PT treatment of hypertrophic scars, keloids and intervertebral disc  
PT disease  
PS Disclosure; Fig 3; 36pp; English.  
CC The original source of the protein material was H-ras transformed  
CC human bronchial epithelial cells (TBE-1). The AA sequence was then  
CC used to develop oligonucleotide probes which were used to screen a  
CC cDNA library of human skin fibroblast mRNA. The longest clone, pGEL  
CC 186.2, represented almost the full gelatinase mRNA sequence except  
CC the leader sequence encoding the first few AA's of the signal  
CC peptide.  
CC Sequence 631 AA;

Query Match 54.8%; Score 34; DB 1; Length 631;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKFKFK 8  
||| |::: 269  
Db 262 CKFFFRFQ 269

## RESULT 13

ID W41112 standard; Protein; 429 AA.  
AC W41112;  
DT 08-JUN-1998 (first entry)

DE Human matrix metalloproteinase huMMP-2 (aa203-631).  
KW Matrix metalloproteinase; MMP-2; huMMP-2; human; angiogenesis;  
KW inhibitor; antagonist; integrin alpha-v beta-3;  
KW vitronectin receptor; rheumatoid arthritis; tumour; metastasis;  
KW diabetic retinopathy; macular degeneration; restenosis; therapy.  
OS Homo sapiens.  
PN W09745137-A1.  
PD 04-DEC-1997.  
PF 30-MAY-1997; U09158.  
PR 31-MAY-1996; US-018733.  
PR 31-MAY-1996; US-015869.  
PA (SCRI) SCRIPPS RES INST.  
PI Brooks P, Cheres DA;  
DR WPI; 98-032334/03.  
PT Packaging material containing polypeptide antagonist of alphav,  
PT beta3 integrin - used for inhibition of angiogenesis, and for  
PT treating tumours, inflammation, eye diseases etc.  
PS Example 4; Page 177-179; 234pp; English.  
CC This polypeptide comprises amino acid residues 203-631 of human  
CC mature matrix metalloproteinase 2 (huMMP-2). It was produced by  
CC recombinant methods involving PCR amplification (see V12509) of  
CC huMMP-2 coding sequence and cloning into e.g. pGEX-11ambdat vector  
CC for expression in E. coli as a glutathione-S-transferase fusion  
CC protein. The invention relates to the discovery that angiogenesis  
CC is mediated by the specific vitronectin receptor alphav beta-3, and  
CC that inhibition of alphav beta-3 function inhibits angiogenesis.  
CC Claimed antagonists of alphav beta-3 comprise C-terminal fragments  
CC (see W41083-94) of human or chicken MMP-2, fusion polypeptides,  
CC cyclic or linear polypeptides (see also W41098-110), derivatised  
CC polypeptides, a monoclonal antibody or organic mimetic compound.  
CC The antagonists are used to inhibit angiogenesis in: inflamed  
CC tissue for treatment of arthritis or rheumatoid arthritis; solid  
CC tumours or metastases, particularly to induce tumour regression or  
CC inhibit growth of tumours; and in ocular disorders such as diabetic  
CC retinopathy or macular degeneration (all claimed). They can also  
CC be used to treat restenosis caused by migration of smooth muscle  
CC cells following angioplasty and to reduce blood supply to selected  
CC tissues (claimed). The new antagonists are highly selective for  
CC angiogenesis. Only new blood vessels express alphav beta-3, so  
CC mature vessels are unaffected, and the antagonists should be of low  
CC toxicity. CC Sequence 429 AA;

Query Match 54.8%; Score 34; DB 1; Length 429;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CKFKFKFK 8  
||| |::: 67  
Db 60 CKFFFRFQ 67

## RESULT 14

ID W41226 standard; protein; 631 AA.  
AC W41226;  
DT 09-JUN-1998 (first entry)  
DE Human mature matrix metalloproteinase-2 (MMP-2) protein sequence.  
KW Matrix metalloproteinase-2; MMP-2; alpha-v-beta-5 antagonist; treatment;  
KW vitronectin receptor; inhibition; angiogenesis; integrin; tumour growth;  
KW restenosis; neovascularisation.  
OS Homo sapiens.  
PN W09745447-A1.  
PD 04-DEC-1997.  
PF 30-MAY-1997; U09099.  
PR 31-MAY-1996; US-018733.  
PR 31-MAY-1996; US-015869.  
PA (SCRI) SCRIPPS RES INST.  
PI Brooks P, Cheres DA, Friedlander M;  
DR WPI; 98-041758/04.  
PT Packaging material containing polypeptide antagonist of alphav,  
PT beta5 integrin - used for inhibition of angiogenesis, and for

PT treating tumours, inflammation, eye diseases etc.  
 FS Disclosure; Fig 16; 117pp; English.

CC The present sequence represents the mature protein of human matrix  
 CC metalloproteinase-2 (MMP-2). Fragments of this protein (W41228-33) are  
 CC able to act as alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a  
 CC vitronectin receptor. Inhibitors of alpha-v-beta-5 can inhibit  
 CC angiogenesis. The specification describes a novel labelled package that  
 CC contains an inhibitor of angiogenesis i.e. an alpha-v-beta-5 antagonising  
 CC polypeptide that binds to integrin alpha-v-beta-5 and includes a part of  
 CC the C-terminal domain of MMP. The antagonists are used to inhibit  
 CC angiogenesis in inflamed tissue, in solid tumours or metastases, and in  
 CC a wide range of ocular disorders (e.g. diabetic or other forms of  
 CC retinopathy, neovascular glaucoma, or corneal transplants). They are  
 CC particularly used to induce regression or to inhibit growth of tumours.  
 CC The alpha-v-beta-5 antagonists can also be used to treat restenosis  
 CC caused by migration of smooth muscle cells following angioplasty and to  
 CC reduce blood supply to selected tissues. The antagonists particularly  
 CC inhibit neovascularisation where this is induced by cytokines,  
 CC e.g. transforming growth factor alpha, epidermal growth factor or  
 CC especially vascular endothelial growth factor.  
 SQ Sequence 631 AA;

Query Match 54.8%; Score 34; DB 1; Length 631;

Best Local Similarity 62.5%; Pred. No. 1.6e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 CKFKFKFK 8

DB 262 CKFFPRFQ 269

# RESULT 15

W98821

ID W98821 standard; Protein; 449 AA.

AC W98821;

DT 31-MAR-1999 (first entry)

DE H. pylori GHPO 1428 protein.

KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

KW peptic ulcer disease.

OS Helicobacter pylori.

PN W09843478-AL.

PD 08-OCT-1998.

PF 01-APR-1998; U06371.

PR 29-JUL-1997; US-902615.

PR 01-APR-1997; US-833457.

PR 24-JUN-1997; US-881227.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

DR WPI; 98-542293/46.

DR N-PSDB; X14540.

PT New isolated Helicobacter polynucleotides - used to develop products

PT for the diagnosis, prevention and treatment of Helicobacter

PT infections and gastrointestinal diseases

PS Claim 8; Page 1804-1806; 2054pp; English.

CC This sequence represents a Helicobacter pylori GHPO protein of the

CC invention. The polypeptides can be used for preventing or treating

CC Helicobacter infections, and gastroduodenal diseases associated with

CC these infections, including acute, chronic, and atrophic gastritis, and

CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be

CC used for the production of antibodies. The products can also be used for

CC detection and diagnosis.

SQ Sequence 449 AA;

Query Match

Best Local Similarity 54.8%; Score 34; DB 1; Length 449;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 KFKFKFKFC 10

DB 138 KFAFEKFC 146

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:17:34 ; Search time 41.46 Seconds  
(without alignments)  
14.844 Million cell updates/sec

Title: US-09-124-280A-44  
Perfect score: 55  
Sequence: 1 KWAKRFLK 10

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL\_10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	69.1	469	3 059710	O59710 schizosacch
2	37	67.3	1955	3 074475	O74475 schizosacch
3	37	67.3	1324	4 015312	Q15312 homo sapien
4	37	67.3	679	5 062137	O62137 caenorhabdi
5	37	67.3	660	5 062138	O62138 caenorhabdi
6	37	67.3	662	5 062139	O62139 caenorhabdi
7	36	65.5	93	1 029641	O29641 archaeglob
8	36	65.5	354	3 012154	Q12154 saccharomyc
9	36	65.5	321	5 016430	O16430 caenorhabdi
10	35	63.6	373	2 067695	O67695 aquifex aeo
11	35	63.6	568	4 013647	Q13647 homo sapien
12	35	63.6	1691	4 092538	Q92538 homo sapien
13	35	63.6	548	5 027465	O27465 caenorhabdi
14	35	63.6	74	7 057983	O57983 trioxys sin
15	35	63.6	589	10 092871	Q92871 arabidopsis
16	35	63.6	2344	12 086114	O86114 rabbit hemo
17	35	63.6	2344	12 089273	O89273 rabbit hemo
18	35	63.6	2344	12 086117	O86117 rabbit hemo
19	35	63.6	1080	12 083895	O83895 ovine adeno
20	35	63.6	2344	12 086119	O86119 rabbit hemo
21	34	61.8	141	1 050529	Q50529 methanobact
22	34	61.8	1732	2 051817	Q51817 porphyronon
23	34	61.8	1732	2 052050	O52050 porphyronon
24	34	61.8	364	2 096966	P96966 porphyronon
25	34	61.8	1723	2 072197	P72197 porphyronon
26	34	61.8	1732	2 007412	O07412 porphyronon
27	34	61.8	292	2 025504	O25504 helicobacte
28	34	61.8	369	2 051304	O51304 borrelia bu
29	34	61.8	131	2 067614	O67614 aquifex aeo

30	34	61.8	456	2 082855	O82855 vibrio para
31	34	61.8	1723	2 072194	P72194 porphyronon
32	34	61.8	156	2 074448	P74448 synchocyst
33	34	61.8	329	5 016431	O16431 caenorhabdi
34	34	61.8	646	5 045076	O45076 caenorhabdi
35	34	61.8	1038	5 061199	O61199 caenorhabdi
36	34	61.8	660	5 017248	Q17248 boophilus m
37	34	61.8	680	5 094196	Q94196 caenorhabdi
38	34	61.8	364	5 002303	O02303 caenorhabdi
39	34	61.8	462	10 004622	O04622 arabidopsis
40	34	61.8	487	10 023384	O23384 arabidopsis
41	33	60.0	337	2 067278	O67278 aquifex aeo
42	33	60.0	4450	2 044528	Q44528 bacillus br
43	33	60.0	2273	3 042823	O42823 saccharomyc
44	33	60.0	490	3 060016	O60016 schizosacch
45	33	60.0	467	11 097431	P97431 mus musculu

## ALIGNMENTS

RESULT 1  
O59710 PRELIMINARY; PRT; 469 AA.  
ID O59710  
AC O59710;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
DE FEREDOXIN-NADP+ REDUCTASE.  
GN SPC3B8.01C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA LYNE M., RAJANDREAM M.A., BARRELL B.G., BECK A., REINHARDT R.,  
RA POHL T;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL022244; CAA18290.1; -  
SQ SEQUENCE 469 AA; 52715 MW; 54915B59 CRC32;

Query Match 69.1%; Score 38; DB 3; Length 469;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KWAKRFLK 10  
Db 402 WSKKFLK 410  
::: |||

RESULT 2  
O74475 PRELIMINARY; PRT; 1955 AA.  
ID O74475  
AC O74475;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE PUTATIVE 1,3-BETA-GLUCAN SYNTHASE COMPONENT.  
GN SPC1840.02C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972H-;  
RA RIEGER M., WOOD V., RAJANDREAM M.A., BARRELL B.G.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL031179; CAA20125.1; -  
SQ SEQUENCE 1955 AA; 225185 MW; C592A33D CRC32;

Query Match 67.3%; Score 37; DB 3; Length 1955;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 DB 1446 WRATKRFK 1454  
 I:| ||| I

RESULT 3  
 Q15312 PRELIMINARY; PRT; 1324 AA.  
 AC Q15312;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE R KAPPA B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA NIEVERS A., BOUMBEESTER T., SCHEIDERREIT C.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92135142.  
 RA ADAMS B.S., LEUNG K.Y., HANLEY E.W., NABEL G.J.;  
 RT "Cloning of R kappa B, a novel DNA-binding protein that recognizes  
 the interleukin-2 receptor alpha chain kappa B site.";  
 RL New Biol. 3:1063-1073(1991).  
 DR EMBL; X80878; CAA56846.1; -. Q -> G (IN REF. 2).  
 FT CONFLICT 97 97  
 SQ SEQUENCE 1324 AA; 141859 MW; E64355AD CRC32;

Query Match 67.3%; Score 37; DB 4; Length 1324;  
 Best Local Similarity 50.0%; Pred. No. 91;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKRFK 10  
 DB 198 EWRTQRYLK 207  
 I:| ||| I

RESULT 4  
 O62137 PRELIMINARY; PRT; 679 AA.  
 AC O62137;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE F08A8.2 PROTEIN.  
 GN F08A8.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.

RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.; III of C.  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99710; CAB16864.1; -.  
 SQ SEQUENCE 679 AA; 77310 MW; E86F3685 CRC32;

Query Match 67.3%; Score 37; DB 5; Length 679;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 DB 514 WKATEKFLK 522  
 I:| ||| I

RESULT 5  
 O62138 PRELIMINARY; PRT; 660 AA.  
 ID O62138;  
 AC O62138;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE F08A8.3 PROTEIN.  
 GN F08A8.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.

RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99710; CAB16865.1; -.  
 SQ SEQUENCE 660 AA; 74781 MW; C24E6FFE CRC32;

Query Match 67.3%; Score 37; DB 5; Length 660;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAQKRFK 10  
 DB 495 WKATEKFLK 503  
 I:| ||| I

RESULT 6  
 O62139 PRELIMINARY; PRT; 662 AA.  
 ID O62139;  
 AC O62139;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE F08A8.4 PROTEIN.  
 GN F08A8.4.



OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL HARRIS B.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans\*;  
 RL Nature 368:32-38(1994).  
 DR EMBL: Z99710; CAB16866.1; -;  
 SQ SEQUENCE 562 AA; 74771 MW; DD0EADA7 CRC32;

Query Match 67.3%; Score 37; DB 5; Length 662;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WKAKQKFLK 10  
 ||| :|||  
 DB 497 WKATEKFLK 505

RESULT 7  
 O29641 ID O29641 PRELIMINARY; PRT; 93 AA.  
 AC O29641;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)  
 DE HYPOTHETICAL 11.1 KD PROTEIN.  
 GN AF0614.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 OC Archaeoglobus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE: 98049343  
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,  
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
 RA OVERBEER R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESSE C.R.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 DR EMBL: AE001062; AAB90629.1; -;  
 DR TIGR: AF0614; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 93 AA; 11148 MW; 0961504C CRC32;

Query Match 65.5%; Score 36; DB 1; Length 93;

Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KWKAQKFLK 10  
 :|| :|||  
 DB 75 EWKSKRFIR 84  
 RESULT 8  
 O12154 ID O12154 PRELIMINARY; PRT; 354 AA.  
 AC O12154;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last annotation update)  
 DE ORF YDL100C FROM LEFT ARM OF CHROMOSOME IV.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY1679;  
 RA BOSKOVIC J., SAIZ J.E., SOLER-MIRA A., GARCIA-CANTALEJO J.,  
 RA REVUELTA J.L., JIMENEZ A., BALLESTA J.P.G., DEL REY F., REMACHA M.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA BALLESTA J.P.G., REMACHA M., SOLER-MIRA A., JIMENEZ A.,  
 RA GARCIA-CANTALEJO J.M., BOSKOVIC J., DEL REY F., REVUELTA J.L.,  
 RA BUITRAGO M.J., SANZ J.E.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X95644; CAA64913.1; -;  
 DR EMBL: Z74148; CAA98667.1; -;  
 SQ SEQUENCE 354 AA; 39353 MW; 63AB1475 CRC32;

Query Match 65.5%; Score 36; DB 3; Length 354;  
 Best Local Similarity 55.6%; Pred. No. 40;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAQKFLK 9  
 :|| :|||  
 DB 291 RWMQKKYL 299

RESULT 9  
 O16430 ID O16430 PRELIMINARY; PRT; 321 AA.  
 AC O16430;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
 DE C02E7.5 PROTEIN.  
 GN C02E7.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA FULTON B., WOHLDMANN P.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF016446; AAC24164.1; -;  
 SQ SEQUENCE 321 AA; 37579 MW; BB77BA41 CRC32;

Query Match 65.5%; Score 36; DB 5; Length 321;  
 Best Local Similarity 66.7%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WKAKRFLK 10  
 |||||  
 Db 232 WKVQQLLK 240

RESULT 10  
 O67695 PRELIMINARY; PRT; 373 AA.  
 AC O67695;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE PEPTIDE CHAIN RELEASE FACTOR RF-2.  
 GN PRFB.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE; 98196666.  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus.";  
 RT Nature 392:353-358(1998).  
 RL Nature 392:353-358(1998).  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE000758; AAC07656.1; -;  
 DR PFAM: PF00472; RF-1.1;  
 SQ SEQUENCE 373 AA; 43309 MW; 921E417E CRC32;

Query Match 63.6%; Score 35; DB 2; Length 373;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQK 6  
 |||||  
 Db 361 KWKAQK 366

RESULT 11  
 Q13647 PRELIMINARY; PRT; 568 AA.  
 AC Q13647;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last annotation update)  
 DE TESTIS SPECIFIC BASIC PROTEIN.

GN TSBP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RA O'HERN P.A., YAVETZ H., MOY T., YAVETZ B., LIANG Z.G., WANG G.Y.,  
 RA GOLDBERG E.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U60665; AAB02975.1; -;  
 SQ SEQUENCE 568 AA; 63319 MW; 2A2CBEE0 CRC32;

Query Match 63.6%; Score 35; DB 4; Length 568;  
 Best Local Similarity 50.0%; Pred. No. 95;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWKAQRFLK 10  
 |||||  
 Db 557 KWKDKKFFK 566

RESULT 12  
 Q92538 PRELIMINARY; PRT; 1691 AA.  
 AC Q92538;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE MYELOBLAST KIAA0248 (FRAGMENT).  
 GN KIAA0248.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BONE MARROW;  
 RX MEDLINE; 97191544.  
 RA NAGASE T., SEKI N., ISHIKAWA K., OHIRA M., KAWARABAYASI Y., OHARA O.,  
 RA TANAKA A., KOTANI H., MIYAJIMA N., NOMURA N.;  
 RT "Prediction of the coding sequences of unidentified human genes. VI.  
 The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by  
 analysis of cDNA clones from cell line KG-1 and brain.";  
 RL DNA Res. 3:321-329(1996).  
 DR EMBL: D87435; BAA13379.1; -;  
 DR PFAM: PF01369; Sec7; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1691 AA; 187806 MW; 8D8EC432 CRC32;

Query Match 63.6%; Score 35; DB 4; Length 1691;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQR 7  
 |||||  
 Db 58 KWKQKR 64

RESULT 13  
 Q27465 PRELIMINARY; PRT; 548 AA.  
 AC Q27465;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE SIMILAR TO CYTOCHROME P450 (EC 1.14.14.1).  
 GN B0304.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1].  
 RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.; III of C.  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA GEISEL C.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; U39472; AAA80134.1; --  
 DR PFAM; PF00067; P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.  
 FT BINDING 471 471 HEME (BY SIMILARITY).  
 SQ SEQUENCE 548 AA; 63665 MW; C4F8339E CRC32;

Query Match 63.6%; Score 35; DB 5; Length 548;  
 Best Local Similarity 62.5%; Pred. No. 92;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KWKAQKRF 8  
 ||:|:|  
 Db 146 KWQRQRF 153

RESULT 14  
 O97983 PRELIMINARY; PRT; 74 AA.  
 AC O97983;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE MHC CLASS I ALPHA 2 CHAIN (FRAGMENT).  
 OS Trionyx sinensis (Chinese softshell turtle) (Pelodiscus sinensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Testudines;  
 OC Cryptodira; Trionychidae; Trionychidae; Trionyx.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA XIA C.;  
 RT "MHC Class I Alpha 2 Domain of Soft-shelled turtle (Trionyx  
 RT sinensis).";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022885; CAB19842.1; --  
 KW MHC.  
 FT NON\_TER 1 1  
 FT NON\_TER 74 74  
 SQ SEQUENCE 74 AA; 8670 MW; FCB19E68 CRC32;

Query Match 63.6%; Score 35; DB 7; Length 74;  
 Best Local Similarity 50.0%; Pred. No. 14;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KWKAQKRELK 10  
 ||:|:|  
 Db 50 KWEAERFLQ 59

RESULT 15  
 ID O92RT1 PRELIMINARY; PRT; 589 AA.  
 AC O92RT1;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE GAMMA RESPONSE I PROTEIN.  
 GN GR I.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA DEVEAUX Y., KAZMAIER M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ131708; CAA10484.1; --  
 SQ SEQUENCE 589 AA; 67472 MW; 798F68B8 CRC32;

Query Match 63.6%; Score 35; DB 10; Length 589;  
 Best Local Similarity 66.7%; Pred. No. 98;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 KWKAQKRF 9  
 ||:|:|  
 Db 331 KWKSQHTFL 339

Search completed: September 7, 1999, 23:17:36  
 Job time: 1779 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 1999, 00:43:02 ; Search time 21.13 Seconds  
(without alignments)  
13.378 Million cell updates/sec

Title: US-09-124-280A-44

Perfect score: 55

Sequence: 1 KWKAQRFLK 10

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	483	1	BPI_HUMAN
2	44	80.0	482	1	BPI_BOVIN
3	41	74.5	481	1	LBP_MOUSE
4	39	70.9	445	1	BPI_RABIT
5	38	69.1	493	1	YDAK_CAEEL
6	37	67.3	1154	1	WCL_NEUCR
7	36	65.5	538	1	CP18_DROME
8	36	65.5	452	1	PLM1_PLAFA
9	36	65.5	140	1	RR28_ARATH
10	36	65.5	264	1	VEXB_SALTI
11	36	65.5	571	1	YHY4_YEAST
12	36	65.5	434	1	YSX4_CAEEL
13	35	63.6	560	1	CBS_RAT
14	35	63.6	1851	1	CCAI_DROME
15	35	63.6	481	1	LBP_RAT
16	35	63.6	1002	1	ODOL_HUMAN
17	35	63.6	2344	1	POLN_RHDV
18	34	61.8	198	1	ATPF_MYCGA
19	34	61.8	550	1	CBS_HUMAN
20	34	61.8	898	1	DPOL_BP74
21	34	61.8	481	1	LBP_HUMAN
22	34	61.8	285	1	YDH5_SCHPO
23	34	61.8	461	1	YUL2_CAEEL
24	33	60.0	883	1	CAPP_ECOLI
25	33	60.0	286	1	CBPX_PEA
26	33	60.0	491	1	CPB1_RAT
27	33	60.0	491	1	CPB2_RAT
28	33	60.0	500	1	CPJ1_RABIT
29	33	60.0	502	1	CPJ2_HUMAN
30	33	60.0	502	1	CPJ3_RAT
31	33	60.0	501	1	CPJ6_MOUSE
32	33	60.0	1895	1	GLS2_YEAST
33	33	60.0	4451	1	GRSB_BACBR
34	33	60.0	2273	1	HPA1_YEAST
35	33	60.0	467	1	IRF6_HUMAN
36	33	60.0	482	1	LBP_RABIT
37	33	60.0	151	1	RR28_TOBAC
38	33	60.0	164	1	RP07_VACCV
39	33	60.0	164	1	RP07_VARV
40	33	60.0	287	1	YD11_MYCQE
41	33	60.0	238	1	Y354_AQAE
42	33	60.0	67	1	YPHE_BACSU
43	32	58.2	105	1	2SS2_CAPMA

#### ALIGNMENTS

##### RESULT 1

ID	BPI_HUMAN	STANDARD;	PRT;	483 AA.	
AC	P17213;				
DT	01-AUG-1990 (REL. 15, CREATED)				
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI) (CAP 57).				
GN	BPI.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 28-64.				
RX	MEDLINE; 89255455.				
RA	GRAY P.W., FLAGGS G., LEONG S.R., GUMINA R.J., WEISS J., OOI C.E.,				
RA	ELSBACH P.;				
RT	"Cloning of the cDNA of a human neutrophil bactericidal protein.				
RT	Structural and functional correlations.";				
RT	J. BIOL. CHEM. 264:9505-9509(1989).				
RN	[2]				
RP	SEQUENCE OF 28-42.				
RX	MEDLINE; 88033057.				
RA	OOI C.E., WEISS J., ELSBACH P., FRANGIONE B., MANNION B.;				
RT	"A 25-kDa NH2-terminal fragment carries all the antibacterial				
RT	activities of the human neutrophil 60-kDa				
RT	bactericidal/permeability-increasing protein.";				
RL	J. BIOL. CHEM. 262:14891-14894(1987).				
RN	[3]				
RP	SEQUENCE OF 28-47.				
RX	MEDLINE; 89315847.				
RA	GABAY J.E., SCOTT R.W., CAMPANELLI D., GRIFFITH J., WILDE C.,				
RA	MARRA M.N., SEEGER M., NATHAN C.F.;				
RT	"Antibiotic proteins of human polymorphonuclear leukocytes.";				
RT	PROC. NATL. ACAD. SCI. U.S.A. 86:5610-5614(1989).				
RN	[4]				
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).				
RX	MEDLINE; 97334442.				
RA	BEAMER L.J., CARROLL S.F., EISENBERG D.;				
RT	"Crystal structure of human BPI and two bound phospholipids at 2.4-A				
RT	resolution.";				
RL	SCIENCE 276:1861-1864(1997).				
CC	-1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES				
CC	OF GRAM-NEGATIVE BACTERIA: THIS SPECIFICITY MAY BE EXPLAINED BY A				
CC	STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE				
CC	NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE				
CC	GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR				
CC	LEUKOCYTES (PMN) GRANULES.				
CC	-1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES.				
CC	-1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF				
CC	THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE				
CC	MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE				
CC	RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-				
CC	AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED				
CC	RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE				
CC	INGESTED.				
CC	-1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				

Q16181 homo sapien  
Q09571 caenorhabdi

```

CC EMBL; J04739; G179529; ALT_INIT.
DR PIR; A29464; A29464.
DR PIR; A30909; A30909.
DR PIR; A33850; A33850.
DR PDB; 1BP1; 04-SEP-97.
DR MIM; 109195; -.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
KW ANTIBIOTIC; SIGNAL; TRANSMEMBRANE; GLYCOPROTEIN; 3D-STRUCTURE.
FT SIGNAL 1 27
FT CHAIN 28 483
FT SITE 236 241
FT TRANSMEM 365 385
FT SEQUENCE 483 AA; 53396 MW; 0AACEF65 CRC32;
SQ

Query Match 100.0%; Score 55; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10
Db 117 KWKAKRFLK 126

RESULT 2
BPI_BOVIN
ID BPI_BOVIN STANDARD; PRT; 482 AA.
AC P17453;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN PRECURSOR (BPI).
GN BPI.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE; 90272418.
RA LEONG S.R., CAMERATO T.;
RT "Nucleotide sequence of the bovine bactericidal permeability
RL NUCLEIC ACIDS RES. 18:3052-3052(1990).
CC -!- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES
CC OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A
CC STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE
CC NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE
CC GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR
CC LEUKOCYTES (PMN) GRANULES (BY SIMILARITY)
CC -!- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY
CC SIMILARITY).
CC -!- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF
CC THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE
CC MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE
CC RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-
CC AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED
CC RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE
CC INGESTED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X99347; E253895;
MGI; 1098776; LBP;
MCD; MGI.1098776; LBP;
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
DR HSSP; P17213; 1BP1.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 24
FT CHAIN 25 481
FT CARBOHYD 300 300
FT CARBOHYD 355 355
FT SEQUENCE 481 AA; 53312 MW; 973CFF58 CRC32;
SQ

```

```

DR EMBL; X52563; G139; -.
DR PIR; S10180; S10180.
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
DR HSSP; P17213; 1BP1.
KW ANTIBIOTIC; SIGNAL; MEMBRANE; GLYCOPROTEIN.
FT SIGNAL 1 26
FT CHAIN 27 482
FT SITE 235 240
FT CARBOHYD 62 62
FT CARBOHYD 303 303
FT CARBOHYD 375 375
FT CARBOHYD 389 389
FT CARBOHYD 463 463
FT SEQUENCE 482 AA; 53432 MW; 20BA9CFA CRC32;
SQ

Query Match 80.0%; Score 44; DB 1; Length 482;
Best Local Similarity 70.0%; Pred. No. 0.53;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10
Db 116 KWKAKRFLK 125

RESULT 3
LBP_MOUSE
ID LBP_MOUSE STANDARD; PRT; 481 AA.
AC Q61805;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).
GN LBP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE; 97289150.
RA LENGACHER S., JONGENEEL C.V., LE ROY D., LEE J.D., KRAVCHENKO V.,
RA ULIVITCH R.J., GLAUSER M.P., HEUMANN D.;
RT "Reactivity of murine and human recombinant LPS-binding protein (LBP)
RT within LPS and gram negative bacteria."
RL J. INFAMM. 47:165-172(1995).
CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
CC TO INTERACT WITH THE CD14 RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X99347; E253895;
MGI; 1098776; LBP;
MCD; MGI.1098776; LBP;
DR PROSITE; PS00400; LBP_BPI_CETP; 1.
DR PFAM; PF01273; LBP_BPI_CETP; 1.
DR HSSP; P17213; 1BP1.
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 24
FT CHAIN 25 481
FT CARBOHYD 300 300
FT CARBOHYD 355 355
FT SEQUENCE 481 AA; 53312 MW; 973CFF58 CRC32;
SQ

```

Query Match 74.5% Score 41; DB 1; Length 481;  
 Best Local Similarity 70.0%; Pred. No. 1.9;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWAKQKFLK 10  
 ||| : |||  
 DB 115 KWVKRKSFLK 124

RESULT 4  
 BPI\_RABBIT STANDARD; PRT; 445 AA.  
 AC Q28739;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE BACTERICIDAL PERMEABILITY-INCREASING PROTEIN (BPI) (FRAGMENT).  
 GN BPI.

OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-NEW ZEALAND WHITE; TISSUE-BONE MARROW;

RA WEISS J., WEINRAUCH Y., LEVY O., FLYNN S.;

RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS

CC -1- FUNCTION: THE CYTOTOXIC ACTION OF BPI IS LIMITED TO MANY SPECIES

CC OF GRAM-NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A

CC STRONG AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE

CC NEGATIVELY CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE

CC GRAM-NEGATIVE BACTERIAL OUTER ENVELOPE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED IN POLYMORPHONUCLEAR

CC LEUKOCYTES (PMN) GRANULES (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: RESTRICTED TO CELLS OF THE MYELOID SERIES (BY

CC SIMILARITY).

CC -1- DOMAIN: THE N-TERMINAL REGION MAY BE EXPOSED TO THE INTERIOR OF

CC THE GRANULE, WHEREAS THE C-TERMINAL PORTION MAY BE EMBEDDED IN THE

CC MEMBRANE. DURING PHAGOCYTOSIS AND DEGRANULATION, PROTEASES MAY BE

CC RELEASED AND ACTIVATED AND CLEAVE BPI AT THE JUNCTION OF THE N-

CC AND C-TERMINAL PORTIONS OF THE MOLECULE, PROVIDING CONTROLLED

CC RELEASE OF THE N-TERMINAL ANTIBACTERIAL FRAGMENT WHEN BACTERIA ARE

CC INGESTED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U61270; G1418284; .

DR PROSITE; PS00400; LBP\_BPI\_CETP; PARTIAL.

DR PFAM; PF01273; LBP\_BPI\_CETP; 1.

DR HSP; P17213; LBPI.

KW ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN.

FT NON\_TER 1

FT SITE 198 203 CLEAVAGE SITES FOR ELASTASE (POTENTIAL).

FT TRANSMEM 327 347 POTENTIAL.

FT CARBOHYD 352 352 POTENTIAL.

SQ SEQUENCE 445 AA; 48837 MW; 5492B04C CRC32;

Query Match 70.9% Score 39; DB 1; Length 445;

Best Local Similarity 60.0%; Pred. No. 4.2;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWAKQKFLK 10  
 :|||:|:|  
 DB 80 RWKARKGFIK 89

RESULT 5  
 YDAK\_CAEL STANDARD; PRT; 493 AA.

AC P90771; O02645;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE PUTATIVE CYTOCHROME P450 C34B7.3 IN CHROMOSOME I (EC 1.14.-.-).

GN C34B7.3

OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;

OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA HARRIS B.;

RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [2]

RP REVISIONS.

RC STRAIN-BRISTOL N2;

RA JONES S.J.M.;

RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE

CC MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED

CC COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; 283220; E1344512; .

DR WORMPEP; C34B7.3; CE08567.

DR PROSITE; PS00086; CYTOCHROME\_P450; 1.

DR PFAM; PF00067; p450; 1.

KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; MONOOXYGENASE; HEME;

KW TRANSMEMBRANE.

FT TRANSMEM 1 21 POTENTIAL.

FT TRANSMEM 60 80 POTENTIAL.

FT TRANSMEM 290 310 POTENTIAL.

FT BINDING 440 440 HEME (BY SIMILARITY).

SQ SEQUENCE 493 AA; 56981 MW; D52DC3DD CRC32;

Query Match 69.1% Score 38; DB 1; Length 493;

Best Local Similarity 85.7%; Pred. No. 7.2;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WKAOKRF 8  
 ||||:|  
 DB 118 WKAORRF 124

RESULT 6  
 WCL\_NEUCR STANDARD; PRT; 1154 AA.

ID WCL\_NEUCR

AC Q01371;

DT 01-NOV-1997 (REL. 35, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE WHITE COLLAR 1 PROTEIN (WCL).

GN WC-1.

OS NEUROSPORA CRASSA.

OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;

OC SORDARIALES; SORDARIACEAE; NEUROSPORA.

RN [1]





CC CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 33-PHE-1-LEU-34 BOND IN THE  
 CC CC ALPHA-CHAIN OF HEMOGLOBIN, LEADING TO DENATURATION OF MOLECULE.  
 CC CC -1- SUBCELLULAR LOCATION: VACUOLAR. COULD BE FIRST ANCHORED TO THE  
 CC CC MEMBRANE THROUGH ITS PROPEPTIDE BEFORE BEING RELEASED.  
 CC CC -1- DEVELOPMENTAL STAGE: ERYTHROCYTIC STAGES.  
 CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE  
 CC CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.  
 CC CC  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL; X75787; G482941; --  
 CC PIR; PT0434; PT0434  
 CC PROSITE; PS00141; ASP\_PROTEASE; 2.  
 CC PFAM; PF00026; asp; 1.  
 CC DR HSSP; P46925; ISME.  
 CC KW HYDROLASE; ASPARTYL PROTEASE; GLYCOPROTEIN; ZYMOGEN; SIGNAL.  
 CC FT SIGNAL 1 ?  
 CC FT PROPEP 1 124  
 CC FT CHAIN 125 452  
 CC FT ACT\_SITE 157 157  
 CC FT ACT\_SITE 337 337  
 CC FT DISULFID 170 175  
 CC FT DISULFID 372 408  
 CC FT CARBOHYD 184 184  
 CC FT CARBOHYD 218 218  
 CC FT CARBOHYD 326 326  
 CC FT CARBOHYD 440 440  
 CC FT CARBOHYD 440 440  
 CC SQ SEQUENCE 452 AA; 51461 MW; AD192ADE CRC32;

Query Match 65.5%; Score 36; DB 1; Length 452;  
 Best Local Similarity 85.7%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 WKAQKRF 8  
 |||||  
 Db 32 WKIQRF 38  
 RESULT 9  
 RK28\_ARATH  
 ID K22795;  
 AC R02795;  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
 DE 50S RIBOSOMAL PROTEIN L28, CHLOROPLAST PRECURSOR.  
 GN RPL28 OR F4P9.22.  
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
 CC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA;  
 CC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 CC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA ROUNSELY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,  
 RA SYKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R.,  
 RA VENTER J.C.;  
 RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC CC  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC

CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL; AC002332; G2459427; --  
 CC PFAM; PF00830; Ribosomal\_L28; 1.  
 CC KW RIBOSOMAL PROTEIN; TRANSIT PEPTIDE; CHLOROPLAST.  
 CC FT TRANSIT 1 63  
 CC FT CHAIN 64 140  
 CC FT CHAIN 64 140  
 CC SQ SEQUENCE 140 AA; 15699 MW; B600C960 CRC32;  
 CC  
 CC Query Match 65.5%; Score 36; DB 1; Length 140;  
 Best Local Similarity 66.7%; Pred. No. 4.6;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 WKAQKRF 10  
 |||||  
 Db 102 WEAGKRFVK 110  
 RESULT 10  
 VEXB\_SALTI  
 ID VEXB\_SALTI STANDARD; PRT; 264 AA.  
 AC P43109;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE VI POLYSACCHARIDE EXPORT INNER-MEMBRANE PROTEIN VEXB.  
 GN VEXB.  
 OS SALMONELLA TYPHI.  
 CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 CC SALMONELLA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GIU 10007;  
 RX MEDLINE; 93322324.  
 RA HASHIMOTO Y., LI N., YOKOYAMA H., EZAKI T.;  
 RT "Complete nucleotide sequence and molecular characterization of Viab  
 RL region encoding Vi antigen in Salmonella typhi.";  
 RL J. BACTERIOL. 175:4456-4465(1993).  
 CC -1- FUNCTION: MAY FORM AN ATP-DRIVEN VI POLYSACCHARIDE EXPORT  
 CC APPARATUS, IN ASSOCIATION WITH THE VEXA, VEXC AND VEXD PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
 CC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE ABC-2 SUBFAMILY OF INTEGRAL MEMBRANE  
 CC PROTEINS.  
 CC CC  
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC CC the European Bioinformatics Institute. There are no restrictions on its  
 CC CC use by non-profit institutions as long as its content is in no way  
 CC CC modified and this statement is not removed. Usage by and for commercial  
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL; D14156; D1003706; --  
 CC PROSITE; PS00890; ABC2\_Membrane; 1.  
 CC PFAM; PF01061; ABC2\_Membrane; 1.  
 CC KW POLYSACCHARIDE TRANSPORT; TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE.  
 CC FT TRANSMEM 38 58  
 CC FT TRANSMEM 58 85  
 CC FT TRANSMEM 65 85  
 CC FT TRANSMEM 120 140  
 CC FT TRANSMEM 149 169  
 CC FT TRANSMEM 191 211  
 CC FT TRANSMEM 239 259  
 CC SQ SEQUENCE 264 AA; 30429 MW; 52811E4C CRC32;  
 CC  
 CC Query Match 65.5%; Score 36; DB 1; Length 264;  
 Best Local Similarity 66.7%; Pred. No. 8.8;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 KWKAQKREL 9  
 :||:| | | |

```
Db 142 RWEAKPFL 150
RESULT 11
YH4_YEAST
ID YH4_YEAST STANDARD; PRT; 571 AA.
AC P38871;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 65.8 KD PROTEIN IN GNDI-IK1 INTERGENIC REGION.
GN YH184W.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;
OC SACHAROMYCETACEAE; SACHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE; 94378003.
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
RA VAUDIN M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL SCIENCE 265:2077-2082(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00028; G458914; -.
DR PIR; S46675; S46675.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 571 AA; 65815 MW; 1A20600D CRC32;
Query Match 65.5%; Score 36; DB 1; Length 571;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 WRAQKFLK 10
II ::::II
Db 198 WKTEKFLK 206
RESULT 12
YSX4_CAEEL
ID YSX4_CAEEL STANDARD; PRT; 434 AA.
AC Q10023;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 50.4 KD PROTEIN T28D9.4 IN CHROMOSOME II.
GN T28D9.4.
OS CAENORHABDITIS ELLEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX WATERSTON R.;
RA WATERSTON R.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC -----
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00028; G458914; -.
DR PIR; S46675; S46675.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 571 AA; 65815 MW; 1A20600D CRC32;
Query Match 65.5%; Score 36; DB 1; Length 571;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 WRAQKFLK 10
II ::::II
Db 198 WKTEKFLK 206
RESULT 12
YSX4_CAEEL
ID YSX4_CAEEL STANDARD; PRT; 434 AA.
AC Q10023;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 50.4 KD PROTEIN T28D9.4 IN CHROMOSOME II.
GN T28D9.4.
OS CAENORHABDITIS ELLEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX WATERSTON R.;
RA WATERSTON R.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC -----
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28738; G861365; -.
DR WORMPEP; T28D9.4; CE02069.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT DOMAIN 151 154 POLY-LEU.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 314 334 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
SQ SEQUENCE 434 AA; 50393 MW; 95639B3F CRC32;
Query Match 65.5%; Score 36; DB 1; Length 434;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 KWAKQKREL 9
:::IIII
Db 3 RWKPKREL 11
RESULT 13
CBS_RAT
ID CBS_RAT STANDARD; PRT; 560 AA.
AC P32232;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CYSTATHIONINE BETA-SYNTHASE (EC 4.2.1.22) (SERINE SULFHYDRASE)
DE (BETA-THIONASE) (HEMOPROTEIN H-450).
GN CBS.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=LIVER.
RX MEDLINE; 92283859.
RA SWAROOP M., BRADLEY K., OHURA T., TAHARA T., ROPER M.D.,
RA ROSENBERG L.E., KRAUS J.P.;
RT "Rat cystathionine beta-synthase. Gene organization and alternative
RT splicing."
RL J. BIOL. CHEM. 267:11455-11461(1992).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22 AND 38-47.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE; 91210211.
RA ISHIHARA S., MOROHASHI K.-I., SADANO H., KAWABATA S.-I., GOTOH O.,
RA OMURA T.;
RT "Molecular cloning and sequence analysis of cDNA coding for rat liver
RT hemoprotein H-450."
RL J. BIOCHEM. 108:899-902(1990).
CC -!- CATALYTIC ACTIVITY: L-SERINE + L-HOMOCYSTEINE = CYSTATHIONINE +
CC H(2)O.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: FIRST STEP IN HOMOCYSTEINE TRANSULFURATION.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, AND BRAIN.
CC -!- PTM: BINDS COVALENTLY TO A HEME GROUP THROUGH A THIOLATE
CC LIGAND.
CC -!- ALTERNATIVE PRODUCTS: FOUR VARIANTS CAN BE FORMED BY ALTERNATIVE
CC SPLICING (TYPE I THROUGH IV). THE SEQUENCE OF TYPE I IS SHOWN.
CC -!- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.
```

CC -!- SIMILARITY: CONTAINS 1 CBS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M88344; G206597; -;  
 DR EMBL; M88346; G206600; -;  
 DR EMBL; D01098; D1001347; -;  
 DR PIR; JX0145; JX0145.  
 DR PIR; A42790; A42790.  
 DR PROSITE; PS00901; CYS\_SYNTHASE; 1.  
 DR PFAM; PF00291; S\_T\_dehydratase; 1.  
 DR PFAM; PF00571; CBS; 1.  
 KW LYSASE; CYSTEINE BIOSYNTHESIS; PYRIDOXAL PHOSPHATE; HEME; CBS DOMAIN;  
 KW ALTERNATIVE SPLICING.  
 FT INIT\_MET 0  
 FT BINDING 115 115 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 FT DOMAIN 411 464 CBS.  
 FT VARSPIC 513 513 S -> Y (IN TYPE III).  
 FT VARSPIC 514 527 MISSING (IN TYPE III).  
 FT VARSPIC 449 477 AILGMVTLGNMLSSLLAGKVRPSDEVCKV ->  
 FT LQSKDICHPTKRHIQAHGLRKYDTEA (IN TYPE  
 FT IV).  
 FT VARSPIC 478 560 MISSING (IN TYPE III).  
 FT CONFLICT 414 414 L -> P (IN REF. 2).  
 SQ SEQUENCE 560 AA; 61323 MW; 6D972105 CRC32;  
 Query Match 63.68; Score 35; DB 1; Length 560;  
 Best Local Similarity 60.08; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 KWKAKQKFLK 10  
 || || || ||  
 Db 385 KWLQKGF 394  
 RESULT 14  
 CCAL\_DROME STANDARD; PRT; 1851 AA.  
 AC P91645;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CALCIUM CHANNEL ALPHA-1 SUBUNIT (CACOPHONY PROTEIN) (NIGHTBLIND A  
 DE PROTEIN) (NO-ON-TRANSIENT B PROTEIN).  
 GN NBA OR DCAIA OR CAC OR NONB.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
 OC DROSOPHILIDAE; DROSOPHILA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97141514.  
 RA SMITH L.A., WANG X.J., PEIXOTO A.A., NEUMANN E.K., HALL L.M.,  
 RA HALL J.C.;  
 RT "A Drosophila calcium channel alpha subunit gene maps to a genetic  
 RL J. NEUROSCI. 16:7868-7879(1996).  
 CC -!- FUNCTION: CALCIUM CHANNEL. WHEN MUTATED IT IS INVOLVED IN THE  
 CC GENERATION OF BEHAVIORAL, PHYSIOLOGICAL, AND LETHAL PHENOTYPES.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- DEVELOPMENTAL STAGE: ITS EXPRESSION PEAKS IN THE FIRST LARVAL  
 CC INSTAR, MIDPUPAL, AND LATE PUPAL STAGES. IN LATE-STAGE EMBRYOS, IT  
 CC IS EXPRESSED PREFERENTIALLY IN THE NERVOUS SYSTEM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U55776; G1737064; -;  
 DR FLYBASE; FBgn0005563; nba.  
 DR PFAM; PF00520; Ion\_trans; 4.  
 KW IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;  
 KW CALCIUM CHANNEL; GLYCOPROTEIN.  
 FT TRANSMEM 37 57 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 108 128 POTENTIAL.  
 FT TRANSMEM 168 188 POTENTIAL.  
 FT TRANSMEM 257 277 POTENTIAL.  
 FT TRANSMEM 291 311 POTENTIAL.  
 FT TRANSMEM 440 460 POTENTIAL.  
 FT TRANSMEM 470 490 POTENTIAL.  
 FT TRANSMEM 511 531 POTENTIAL.  
 FT TRANSMEM 566 586 POTENTIAL.  
 FT TRANSMEM 607 627 POTENTIAL.  
 FT TRANSMEM 647 667 POTENTIAL.  
 FT TRANSMEM 768 788 POTENTIAL.  
 FT TRANSMEM 808 828 POTENTIAL.  
 FT TRANSMEM 837 857 POTENTIAL.  
 FT TRANSMEM 906 926 POTENTIAL.  
 FT TRANSMEM 975 995 POTENTIAL.  
 FT TRANSMEM 1021 1041 POTENTIAL.  
 FT TRANSMEM 1091 1111 POTENTIAL.  
 FT TRANSMEM 1130 1150 POTENTIAL.  
 FT TRANSMEM 1161 1181 POTENTIAL.  
 FT TRANSMEM 1223 1243 POTENTIAL.  
 FT TRANSMEM 1271 1291 POTENTIAL.  
 FT TRANSMEM 1317 1337 POTENTIAL.  
 FT CARBOHYD 234 234 POTENTIAL.  
 FT CARBOHYD 235 235 POTENTIAL.  
 FT CARBOHYD 865 865 POTENTIAL.  
 FT CARBOHYD 1430 1430 POTENTIAL.  
 SQ SEQUENCE 1851 AA; 212168 MW; 601976A1 CRC32;  
 Query Match 63.68; Score 35; DB 1; Length 1851;  
 Best Local Similarity 71.48; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 WKAOKRF 8  
 || || || ||  
 Db 424 WRAEKRF 430  
 RESULT 15  
 LBP\_RAT STANDARD; PRT; 481 AA.  
 ID LBP\_RAT  
 AC Q63313;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR (LBP).  
 GN LBP.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-LIVER;  
 RX MEDLINE; 94292804.  
 RA SU G.L., FREESWICK P.D., GELLER D.A., WANG Q., SHAPIRO R.A., WAN Y.H.,  
 RA BILLIAR T.R., TWEARDY D.J., SIMMONS R.L., WANG S.C.;  
 RT "Molecular cloning, characterization, and tissue distribution of rat  
 RT lipopolysaccharide binding protein. Evidence for extrahepatic  
 RT expression".  
 RL J. IMMUNOL. 153:743-752(1994).

CC -!- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL  
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER  
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS  
CC TO INTERACT WITH THE CD14 RECEPTOR.  
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L32132; G510715; -  
CC PROSITE: PS00400; LBP\_BPI\_CETP; 1.  
DR PFAM: PF01273; LBP\_BPI\_CETP; 1.  
DR HSP: P17213; LBP1.  
KW LIPID TRANSPORT; ANTIBIOTIC; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 481 LIPOPOLYSACCHARIDE-BINDING PROTEIN.  
FT CARBOHYD 300 300 POTENTIAL.  
FT CARBOHYD 355 355 POTENTIAL.  
SQ SEQUENCE 481 AA; 53600 MW; 0BF55E7D CRC32;

Query Match 63.6%; Score 35; DB 1; Length 481;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 KWKAKRFLK 10  
||| ::|:  
Db 115 KWKRRSEVK 124

Search completed: September 8, 1999, 00:43:02  
Job time: 2617 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:54:52 ; Search time 31.14 Seconds  
(without alignments)  
12.866 Million cell updates/sec

Title: US-09-124-280A-44

Perfect score: 55

Sequence: 1 KWKAKRFLK 10

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	487	2 A30909	bactericidal/permeability-increasing protein precursor - human
2	55	100.0	250	4 S43383	bactericidal/permeability-increasing protein precursor - human
3	44	80.0	482	2 S10180	bactericidal/permeability-increasing protein precursor - human
4	37	67.3	1154	2 S69206	regulator protein
5	37	67.3	1324	2 S52863	DNA-binding protein
6	36	65.5	354	2 S67642	probable arsenical
7	36	65.5	264	2 G36892	V1 polysaccharide
8	36	65.5	93	2 F69326	hypothetical prote
9	36	65.5	571	2 S46675	hypothetical prote
10	36	65.5	452	2 S41717	aspartic hemoglobi
11	35	63.6	2344	1 RRMWRH	genome polyprotein
12	35	63.6	1003	2 A38234	oxoglutarate dehyd
13	35	63.6	2344	2 S55399	genome polyprotein
14	35	63.6	2344	2 S64740	genome polyprotein
15	35	63.6	561	2 A42790	cystathionine beta
16	35	63.6	478	2 C42790	cystathionine beta
17	35	63.6	546	2 B42790	cystathionine beta
18	35	63.6	356	2 D42790	cystathionine beta
19	35	63.6	547	2 JX0145	hemoprotein H-450
20	35	63.6	481	2 I56246	lipopolysaccharide
21	35	63.6	373	2 E70458	translation releas
22	35	63.6	568	2 G02753	testis specific ba
23	35	63.6	1080	2 PC4394	DNA-directed DNA p
24	34	61.8	898	1 DJBPT4	DNA-directed DNA p
25	34	61.8	487	2 C71417	cytochrome P450 -
26	34	61.8	552	2 A55760	cystathionine beta
27	34	61.8	481	2 A54136	lipopolysaccharide
28	34	61.8	477	2 A35843	lipopolysaccharide
29	34	61.8	198	2 S24335	H+-transporting AT
30	34	61.8	369	2 D70140	hypothetical prote
31	34	61.8	292	2 A64624	hypothetical prote
32	34	61.8	156	2 S76420	hypothetical prote
33	34	61.8	462	2 T01732	UTP-glucose glucos
34	34	61.8	131	2 D70448	hypothetical prote
35	33	60.0	491	1 O4RTPB	cytochrome P450 2B
36	33	60.0	491	1 O4RTPB	cytochrome P450 2B
37	33	60.0	164	1 RNVT219	DNA-directed RNA p
38	33	60.0	883	1 QYEC	phosphoenolpyruvat
39	33	60.0	4452	1 YGBSG2	gramicidin S synth

40 33 60.0 151 1 R5NT28 ribosomal protein  
41 33 60.0 501 2 A40938 cytochrome P450 1b  
42 33 60.0 367 2 S19172 cytochrome P450 2B  
43 33 60.0 164 2 F36848 A5R protein - vari  
44 33 60.0 286 2 S62370 serine-type carbox  
45 33 60.0 1704 2 S71363 probable transport

ALIGNMENTS

RESULT 1  
A30909  
bactericidal/permeability-increasing protein precursor - human  
N:Alternate names: 55K bactericidal protein  
C:Species: Homo sapiens (man)  
C:Date: 18-Apr-1989 #sequence\_revision 18-Apr-1989 #text\_change 17-Mar-1999  
C:Accession: A33850; B54136; A29464; A43600; A49716; A30909  
R:Gray, P.W.; Flagg, G.; Leong, S.R.; Gumina, R.J.; Weiss, J.; Ooi, C.E.; Elsbach, P.  
J. Biol. Chem. 264, 9505-9509, 1989  
A:Title: Cloning of the cDNA of a human neutrophil bactericidal protein. Structural a  
A:Reference number: A33850; MUID:89255455  
A:Accession: A33850  
A:Molecule type: mRNA  
A:Residues: 1-487 <GRA>  
A:Cross-references: GB:J04739; NID:gl79528; PID:gl79529  
R:Wilde, C.G.; Seilhamer, J.J.; McGrogan, M.; Ashton, N.; Snable, J.L.; Lane, J.C.; L  
J. Biol. Chem. 269, 17411-17416, 1994  
A:Title: Bactericidal/permeability-increasing protein and lipopolysaccharide (LPS)-bi  
A:Reference number: A54136; MUID:94292492  
A:Accession: B54136  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-374, 'L', 376-487 <WIL>  
A:Experimental source: HL-60 cells  
A:Note: Sequence extracted from NCBI backbone (NCBI:149855)  
R:Ooi, C.E.; Weiss, J.; Elsbach, P.; Frangione, B.; Mannion, B.  
J. Biol. Chem. 262, 14891-14894, 1987  
A:Title: A 25-kDa amino-terminal fragment carries all the antibacterial activities of  
A:Reference number: A29464; MUID:88033057  
A:Accession: A29464  
A:Molecule type: protein  
A:Residues: 32-51 <OOI>  
A:Experimental source: neutrophils  
R:Wasiluk, K.R.; Skubitz, K.M.; Gray, B.H.  
Infect. Immun. 59, 4193-4200, 1991  
A:Title: Comparison of granule proteins from human polymorphonuclear leukocytes which  
A:Reference number: A43600; MUID:92040097  
A:Accession: A43600  
A:Molecule type: protein  
A:Residues: 32-52, 'R' <WAS>  
R:Little, R.G.; Kellner, D.N.; Lim, E.; Burke, D.J.; Conlon, P.J.  
J. Biol. Chem. 269, 1865-1872, 1994  
A:Title: Functional domains of recombinant bactericidal/permeability increasing prote  
A:Reference number: A49716; MUID:94124531  
A:Accession: A49716  
A:Molecule type: protein  
A:Residues: 32-130;132-141;143-165;202-215, 'E', 217-225 <LIT>  
C:Comment: The bactericidal/permeability-increasing protein (BPI) is a 60 kD membrane  
which is specific for gram-negative bacteria. BPI has a high affinity for lipopolysac  
between BPI and an LPS-binding protein from liver and cholesterol ester transfer prot  
C:Genetics:  
A:Gene: GDB:BPI  
A:Cross-references: GDB:131572; OMIM:109195  
A:Map position: 20q11.23-20q12  
C:Superfamily: lipopolysaccharide-binding protein  
C:Keywords: antibacterial; cytotoxin; glycoprotein; heparin binding; neutrophil  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-487/Product: bactericidal permeability-increasing protein #status predicted <MAT  
F:32-51/Region: bactericidal #status predicted  
F:380/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 55; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.0062;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
|||||  
Db 121 KWKAQKRELK 130

RESULT 2

bactericidal/permeability-increasing protein - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*  
C:Date: 20-Oct-1994 #sequence\_revision 15-Feb-1996 #text\_change 15-Feb-1996  
C:Accession: S43383  
R:Qi, S.Y.; Li, Y.; O'Connor, C.D.  
Biochem. J. 296, 711-718, 1994  
A:Title: The region around residue 115 of human bactericidal/permeability-increasing protein of a gene coding for the active domain and characterization of recombinant proteins.  
A:Reference number: S43383  
A:Accession: S43383  
A:Molecule type: DNA  
A:Residues: 1-250 <QIS>

Query Match 100.0%; Score 55; DB 4; Length 250;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
|||||  
Db 91 KWKAQKRELK 100

RESULT 3

S10180  
bactericidal/permeability-increasing protein precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Sep-1997  
C:Accession: S10180  
R:Leong, S.R.; Camerato, T.  
Nucleic Acids Res. 18, 3052, 1990  
A:Title: Nucleotide sequence of the bovine bactericidal permeability increasing protein  
A:Reference number: S10180; MUID:90272418  
A:Accession: S10180  
A:Molecule type: mRNA  
A:Residues: 1-482 <LEO>  
A:Cross-references: EMBL:X52563; NID:gl38; PID:gl39  
C:Superfamily: lipopolysaccharide-binding protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-482/Product: bactericidal permeability increasing protein #status predicted <MAT>

Query Match 80.0%; Score 44; DB 2; Length 482;  
Best Local Similarity 70.0%; Pred. No. 0.73;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
|||||  
Db 116 KWKAQKRELK 125

RESULT 4

S69206  
regulator protein white collar 1 - *Neurospora crassa*  
C:Species: *Neurospora crassa*  
C:Date: 21-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Oct-1997  
C:Accession: S69206  
R:Ballarín, P.; Vittorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.  
EMBO J. 15, 1650-1657, 1996  
A:Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is a  
A:Reference number: S69206; MUID:96203083

Query Match 67.3%; Score 37; DB 2; Length 1154;  
Best Local Similarity 85.7%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WKAKRKF 8  
|||||  
Db 1138 WKAKRKF 1144

RESULT 5

S52863  
DNA-binding protein R kappa B - human  
C:Species: Homo sapiens (man)  
C:Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 10-Sep-1997  
C:Accession: S52863  
R:Nieters, A.; Bouwmeester, T.; Scheidereit, C.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: S52863  
A:Accession: S52863  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1324 <NIE>  
A:Cross-references: EMBL:X80878; NID:g695578; PID:g695579

Query Match 67.3%; Score 37; DB 2; Length 1324;  
Best Local Similarity 50.0%; Pred. No. 44;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
||:|:|:  
Db 198 EWRTQRYLK 207

RESULT 6

S67642  
probable arsenical pump-driving ATPase (EC 3.6.1.-) YDL100c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein D2371  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 18-Sep-1998  
C:Accession: S67642; S67419  
R:Ballesta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.; submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67629  
A:Accession: S67642  
A:Molecule type: DNA  
A:Residues: 1-354 <BAL>  
A:Cross-references: EMBL:Z74148; NID:gl431137; PID:e253031; PID:gl431138; MIPS:YDL100  
A:Experimental source: strain S288C  
R:Boskovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jim submitted to the EMBL Data Library, February 1996  
A:Reference number: S67406  
A:Accession: S67419  
A:Molecule type: DNA  
A:Residues: 1-354 <BOS>  
A:Cross-references: EMBL:X95644; NID:gl199535; PID:e223235; PID:gl199549  
C:Genetics:  
A:Map position: 4L  
A:Note: YDL100c  
C:Superfamily: arsenical pump-driving ATPase  
C:Keywords: ATP; hydrolase; P-loop  
F:25-32/Region: nucleotide-binding motif A (P-loop)

```
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-Dec-1997
C:Accession: S46675
R:Macri, C.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of S. cerevisiae cosmid 9186.
A:Reference number: S46671
A:Accession: S46675
A:Molecule type: DNA
A:Residues: 1-571 <MAC>
A:Cross-references: EMBL:U00028; NID:g458909; PID:g458914; MIPS:YHR184w
C:Genetics:
A:Map position: 8R

Query Match      65.5%; Score 36; DB 2; Length 354;
Best Local Similarity 55.6%; Pred. No. 17;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 9
Db 291 RWRKQKRYL 299

RESULT 7
G36892
Vi polysaccharide capsule transporter vexB - Salmonella typhi
C:Species: Salmonella typhi
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Sep-1997
C:Accession: G36892
R:Hashimoto, Y.; Li, N.; Yokoyama, H.; Ezaki, T.
J. Bacteriol. 175, 4456-4465, 1993
A:Title: Complete nucleotide sequence and molecular characterization of Viab region enc
A:Reference number: A36892; MUID:93322324
A:Accession: G36892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <HAS>
A:Cross-references: GB:D14156; NID:g426443; PID:d1003706; PID:g426450
C:Genetics:
A:Gene: vexB

Query Match      65.5%; Score 36; DB 2; Length 264;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 9
Db 142 RWEAKRFL 150

RESULT 8
F69326
hypothetical protein AF0614 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Jun-1998
C:Accession: F69326
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: F69326
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-93 <KLE>
A:Cross-references: GB:AE001062; GB:AE000782; NID:g2689385; PID:g2650010; TIGR:AF0614

Query Match      65.5%; Score 36; DB 2; Length 93;
Best Local Similarity 50.0%; Pred. No. 4.2;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KWKAQKRF 10
Db 75 EWRKSKRFIR 84

RESULT 9
S46675
hypothetical protein YHR184w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein H9186.1
```

```
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-Dec-1997
C:Accession: S46675
R:Macri, C.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of S. cerevisiae cosmid 9186.
A:Reference number: S46671
A:Accession: S46675
A:Molecule type: DNA
A:Residues: 1-571 <MAC>
A:Cross-references: EMBL:U00028; NID:g458909; PID:g458914; MIPS:YHR184w
C:Genetics:
A:Map position: 8R

Query Match      65.5%; Score 36; DB 2; Length 571;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KWKAQKRF 10
Db 198 WKTEKFLK 206

RESULT 10
S41717
aspartic hemoglobinase (EC 3.4.23.-) - Plasmodium falciparum
C:Species: Plasmodium falciparum
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Jan-1999
C:Accession: S41717; PT0434
R:Francis, S.E.; Gluzman, I.Y.; Oksman, A.; Knickerbocker, A.; Mueller, R.; Bryant, M
EMBO J. 13, 306-317, 1994
A:Title: Molecular characterization and inhibition of a Plasmodium falciparum aspartic
A:Reference number: S41717; MUID:94147975
A:Accession: S41717
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-452 <FRA>
A:Cross-references: GB:X75787; NID:g482940; PID:g482941
R:Goldberg, D.E.; Slater, A.F.G.; Beavis, R.; Chait, B.; Cerami, A.; Henderson, G.B.
J. Exp. Med. 173, 961-969, 1991
A:Title: Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum:
A:Reference number: PT0434; MUID:91178457
A:Accession: PT0434
A:Molecule type: protein
A:Residues: 125-146 <GOI>
A:Note: this sequence reveals 9 of 22 residues identical to the most specific mammal
C:Comment: This enzyme plays a key role in the hemoglobinolytic pathway.
C:Keywords: aspartic proteinase; hydrolase

Query Match      65.5%; Score 36; DB 2; Length 452;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KWKAQKRF 8
Db 32 WKIQKRF 38

RESULT 11
RRWNRH
genome polypotein - rabbit hemorrhagic disease virus
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: rabbit hemorrhagic disease virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C:Accession: A41039
R:Meyers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencin
A:Reference number: A41039; MUID:91361557
A:Accession: A41039
A:Molecule type: genomic RNA
```

A:Residues: 1-2344 <MEY>  
 A:Cross-references: GB:M67473  
 C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase

Query Match 63.6%; Score 35; DB 1; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0;

QY 1 KWKAKRFLK 10  
 : : : |||||

DB 109 RWRRSRFLK 118

## RESULT 12

A38234  
 oxoglutarate dehydrogenase (lipoamide) (EC 1.2.4.2) precursor - human  
 N:Alternate names: 2-oxoglutarate:lipoamide 2-oxidoreductase; alpha-Ketoglutarate dehydrogenase  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 13-Nov-1998  
 C:Accession: A38234  
 R:Koike, K.; Urata, Y.; Goto, S.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 1963-1967, 1992  
 A:Title: Cloning and nucleotide sequence of the cDNA encoding human 2-oxoglutarate dehydrogenase  
 A:Reference number: A38234; MUID:92179301  
 A:Accession: A38234  
 A:Molecule type: mRNA  
 A:Residues: 1-1003 <KO>  
 A:Cross-references: GB:D10523  
 A:Experimental source: fetal liver  
 A:Note: sequence extracted from NCBI backbone (NCBIP:87352)  
 C:Superfamily: oxoglutarate dehydrogenase (lipoamide); thiamine pyrophosphate-binding domain  
 C:Keywords: mitochondrion; oxidoreductase; thiamin pyrophosphate; tricarboxylic acid cycle  
 F:1-40/Domain: transit peptide (mitochondrion) #status predicted <TN>  
 F:41-1003/Product: oxoglutarate dehydrogenase (lipoamide) #status predicted <MAT>  
 F:408-453/Domain: thiamine pyrophosphate-binding domain homology <TPB>

Query Match 63.6%; Score 35; DB 2; Length 1003;  
 Best Local Similarity 62.5%; Pred. No. 79;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KWKAKRFF 8  
 || :|||

DB 271 KWSSEKRF 278

## RESULT 13

S55399  
 genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
 C:Species: rabbit hemorrhagic disease virus  
 A:Variety: isolate BS89  
 C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 08-Sep-1997  
 C:Accession: S55399  
 R:Rossi, C.  
 submitted to the EMBL Data Library, May 1995

A:Reference number: S55399  
 A:Accession: S55399  
 A:Molecule type: genomic RNA  
 A:Residues: 1-2344 <ROS>

A:Cross-references: EMBL:X87607; NID:g854640; PID:g854641  
 A:Experimental source: isolate BS89

C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: polyprotein

Query Match 63.6%; Score 35; DB 2; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10  
 : : : |||||

DB 109 RWRRSRFLK 118

## RESULT 14

S64740

genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
 N:Contains: VP60 protein  
 C:Species: rabbit hemorrhagic disease virus  
 A:Variety: isolate AST/89  
 C:Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 17-Mar-1999  
 C:Accession: S64740; S46944; S49018; S65012  
 R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
 submitted to the EMBL Data Library, May 1995  
 A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by

A:Reference number: S64740

A:Accession: S64740

A:Molecule type: genomic RNA

A:Residues: 1-2344 <CAS>

A:Cross-references: EMBL:Z49271; NID:g1182032; PID:el58171; PID:g1150552

A:Experimental source: isolate AST/89

R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra

submitted to the EMBL Data Library, July 1993

A:Description: Molecular cloning, sequence and expression of the capsid protein gene

A:Reference number: S46944

A:Accession: S46944

A:Molecule type: genomic RNA

A:Residues: 1650-2344 <BOG>

A:Cross-references: EMBL:Z24757; NID:g515622; PID:g515623

A:Experimental source: isolate AST/89

R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.

Virus Res. 27, 219-228, 1993

A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus su

A:Reference number: S49018; MUID:93255896

A:Accession: S49018

A:Molecule type: genomic RNA

A:Residues: 1650-1796 <PAR>

A:Cross-references: EMBL:Z24757

A:Experimental source: isolate AST/89

A:Accession: S65012

A:Molecule type: protein

A:Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>

C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C:Keywords: blocked amino end; polyprotein

Query Match 63.6%; Score 35; DB 2; Length 2344;  
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10

: : : |||||

DB 109 RWRRSRFLK 118

## RESULT 15

A42790

cystathionine beta-synthase (EC 4.2.1.22), splice form I - rat

N:Alternate names: beta-thionase; methylcysteine synthase; serine sulfhydrase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 15-Jan-1999

C:Accession: A42790

R:Swaroop, M.; Bradley, K.; Ohura, T.; Tahara, T.; Roper, M.D.; Rosenberg, L.E.; Krau

J. Biol. Chem. 267, 11455-11461, 1992

A:Title: Rat cystathionine beta-synthase. Gene organization and alternative splicing.

A:Reference number: A42790; MUID:92283859

A:Accession: A42790

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-561 <SWA>

A:Cross-references: GB:M88344; NID:g1364273; PID:g206597

A:Experimental source: strain Sprague-Dawley; liver

A:Note: sequence extracted from NCBI backbone (NCBIP:104763)

C:Superfamily: cystathionine beta-synthase; CBS homology



C;Keywords: alternative splicing; carbon-oxygen lyase; hydro-lyase  
F:417-465/Domain: CBS homology <CBS2>

Query Match 63.6%; Score 35; DB 2; Length 561;  
Best Local Similarity 60.0%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KWKAKRRFLK 10  
|| || ||  
Db 386 KWLQKGFMK 395

Search completed: September 7, 1999, 23:54:54  
Job time: 1905 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:23:08 ; Search time 25.2 Seconds  
(without alignments)  
3.916 Million cell updates/sec

Title: US-09-124-280A-44

Perfect score: 55

Sequence: 1 KWKAQKRFK 10

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/PTUS9\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	487	1 US-08-030-644-2	Sequence 2, Appli
2	55	100.0	487	1 US-08-013-801-2	Sequence 2, Appli
3	55	100.0	487	1 US-08-072-063-2	Sequence 2, Appli
4	55	100.0	487	1 US-08-212-132-2	Sequence 2, Appli
5	55	100.0	487	1 US-08-414-924-2	Sequence 2, Appli
6	55	100.0	35	1 US-08-311-611A-6	Sequence 6, Appli
7	55	100.0	15	1 US-08-311-611A-7	Sequence 7, Appli
8	55	100.0	10	1 US-08-311-611A-8	Sequence 8, Appli
9	55	100.0	16	1 US-08-311-611A-9	Sequence 9, Appli
10	55	100.0	17	1 US-08-311-611A-10	Sequence 10, Appli
11	55	100.0	27	1 US-08-311-611A-11	Sequence 11, Appli
12	55	100.0	15	1 US-08-311-611A-16	Sequence 16, Appli
13	55	100.0	15	1 US-08-311-611A-17	Sequence 17, Appli
14	55	100.0	15	1 US-08-311-611A-18	Sequence 18, Appli
15	55	100.0	15	1 US-08-311-611A-19	Sequence 19, Appli
16	55	100.0	15	1 US-08-311-611A-20	Sequence 20, Appli
17	55	100.0	15	1 US-08-311-611A-49	Sequence 49, Appli
18	55	100.0	15	1 US-08-311-611A-51	Sequence 51, Appli
19	55	100.0	24	1 US-08-311-611A-52	Sequence 52, Appli
20	55	100.0	29	1 US-08-311-611A-53	Sequence 53, Appli
21	55	100.0	20	1 US-08-311-611A-54	Sequence 54, Appli
22	55	100.0	25	1 US-08-311-611A-55	Sequence 55, Appli
23	55	100.0	20	1 US-08-311-611A-57	Sequence 57, Appli
24	55	100.0	20	1 US-08-311-611A-58	Sequence 58, Appli
25	55	100.0	26	1 US-08-311-611A-55	Sequence 65, Appli
26	55	100.0	17	1 US-08-311-611A-66	Sequence 66, Appli
27	55	100.0	17	1 US-08-311-611A-68	Sequence 68, Appli
28	55	100.0	487	1 US-08-311-611A-69	Sequence 69, Appli
29	55	100.0	24	1 US-08-311-611A-70	Sequence 70, Appli
30	55	100.0	487	1 US-08-311-611A-146	Sequence 146, App
31	55	100.0	30	1 US-08-311-611A-149	Sequence 149, App
32	55	100.0	20	1 US-08-311-611A-150	Sequence 150, App
33	55	100.0	20	1 US-08-311-611A-151	Sequence 151, App
34	55	100.0	14	1 US-08-311-611A-164	Sequence 164, App
35	55	100.0	12	1 US-08-311-611A-227	Sequence 227, App
36	55	100.0	487	1 US-08-173-968-2	Sequence 2, Appli
37	55	100.0	221	1 US-08-173-968-3	Sequence 3, Appli
38	55	100.0	487	1 US-08-232-527-2	Sequence 2, Appli
39	55	100.0	35	1 US-08-372-783-6	Sequence 6, Appli

Sequence 7, Appli  
Sequence 8, Appli  
Sequence 9, Appli  
Sequence 10, Appli  
Sequence 11, Appli  
Sequence 4, Appli

#### ALIGNMENTS

##### RESULT 1

US-08-030-644-2

; Sequence 2, Application US/08030644

; Patent No. 5348942

; GENERAL INFORMATION:

; APPLICANT: Little, Roger G. II

; APPLICANT: Gazzano-Santoro, Helene

; APPLICANT: Parent, James Brian

; TITLE OF INVENTION: Therapeutic uses of

; TITLE OF INVENTION: Bactericidal/Permeability-Increasing Protein Products

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60608-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/030,644

; FILING DATE: 19930312

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharp, Jeffrey S.

; REGISTRATION NUMBER: 31,879

; REFERENCE/DOCKET NUMBER: 31229

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 487 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-030-644-2

Query Match

Best Local Similarity 100.0%; Score 55; DB 1; Length 487;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KWKAQKRFK 10

Db 121 KWKAQKRFK 130

##### RESULT 2

US-08-013-801-2

; Sequence 2, Application US/08013801

; Patent No. 5420019

; GENERAL INFORMATION:

; APPLICANT: Theofan, Georgia

; APPLICANT: Horwitz, Arnold

; APPLICANT: Burke, David

; APPLICANT: Baltain, Manik

APPLICANT: Grinna, Lynn S  
TITLE OF INVENTION: Stable Bactericidal/Permeability-  
Increasing Protein Products and Pharmaceutical  
TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical  
TITLE OF INVENTION: Compositions Containing the Same  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
OPERATING SYSTEM: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/013,801  
FILING DATE: 02 FEB 1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Thomas C.  
REGISTRATION NUMBER: P-36,989  
REFERENCE/DOCKET NUMBER: 27129/30911  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/346-5750  
TELEFAX: 312/346-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-013-801-2

Query Match 100.0%; Score 55; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFLK 10  
Db 121 KWKAQRFLK 130

RESULT 3  
US-08-072-063-2  
Sequence 2, Application US/08072063  
Patent No. 5439807  
GENERAL INFORMATION:  
APPLICANT: Theofan, Georgia  
APPLICANT: Grinna, Lynn S  
APPLICANT: Horwitz, Arnold  
TITLE OF INVENTION: BPI-Immunoglobulin Fusion Proteins  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
OPERATING SYSTEM: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/072,063  
FILING DATE: 19930519

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers Thomas C.  
REGISTRATION NUMBER: 36,989  
REFERENCE/DOCKET NUMBER: 30659  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-072-063-2

Query Match 100.0%; Score 55; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQRFLK 10  
Db 121 KWKAQRFLK 130

RESULT 4  
US-08-212-132-2  
Sequence 2, Application US/08212132  
Patent No. 5447913  
GENERAL INFORMATION:  
APPLICANT: Little, Roger G.  
APPLICANT: Ammons, William Steve  
TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-  
Increasing Protein Dimer Products  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,132  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 27129/31735  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-212-132-2

Query Match 100.0%; Score 55; DB 1; Length 487;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KWKAQKRFK 10
Db 121 KWKAQKRFK 130

RESULT 5
US-08-414-924-2
; Sequence 2, Application US/08414924
; Patent No. 5494896
; GENERAL INFORMATION:
; APPLICANT: Hansbrough, John F.
; TITLE OF INVENTION: Method of Treating Conditions
; TITLE OF INVENTION: Associated With Burn Injuries
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,924
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/32297
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-414-924-2

```

```

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KWKAQKRFK 10
Db 121 KWKAQKRFK 130

RESULT 6
US-08-311-611A-6
; Sequence 6, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

```

```

; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "Domain II"
US-08-311-611A-6

```

```

Query Match 100.0%; Score 55; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KWKAQKRFK 10
Db 26 KWKAQKRFK 35

```

```

RESULT 7
US-08-311-611A-7
; Sequence 7, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/311,611A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/273,401  
;; FILING DATE: 11-JUL-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/125,651  
;; FILING DATE: 22-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 15 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.2"  
US-08-311-611A-7

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10  
Db 6 KWKAKRFLK 15

RESULT 8  
US-08-311-611A-8  
; Sequence 8, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:

;; NAME: Sharp, Jeffrey S.  
;; REGISTRATION NUMBER: 31,879  
;; REFERENCE/DOCKET NUMBER: 32251  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312/474-6300  
;; TELEFAX: 312/474-0448  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; OTHER INFORMATION: "BPI.8"  
US-08-311-611A-8

Query Match 100.0%; Score 55; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00091;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKRFLK 10  
Db 1 KWKAKRFLK 10

RESULT 9  
US-08-311-611A-9  
; Sequence 9, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids

;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc:feature  
;; OTHER INFORMATION: "BPI.58"  
US-08-311-611A-9

Query Match 100.0%; Score 55; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
Db 7 KWKAQKRFK 16

RESULT 10  
US-08-311-611A-10  
; Sequence 10, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc:feature  
; OTHER INFORMATION: "BPI.65 oxidized"  
US-08-311-611A-10

Query Match 100.0%; Score 55; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KWKAQKRFK 10  
Db 7 KWKAQKRFK 16

RESULT 11  
US-08-311-611A-11  
; Sequence 11, Application US/08311611A  
; Patent No. 5523288  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Jonathan  
; APPLICANT: Kung, Ada H.C.  
; APPLICANT: Lambert, Jr., Lewis H.  
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
; TITLE OF INVENTION: Infection by Administration of  
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
; TITLE OF INVENTION:  
; NUMBER OF SEQUENCES: 227  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311.611A  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/273,401  
; FILING DATE: 11-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/125,651  
; FILING DATE: 22-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharp, Jeffrey S.  
; REGISTRATION NUMBER: 31,879  
; REFERENCE/DOCKET NUMBER: 32251  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: misc:feature  
; OTHER INFORMATION: "BPI.13"  
US-08-311-611A-11

Query Match 100.0%; Score 55; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRFK 10  
Db 18 KWKAQKRFK 27

RESULT 12  
US-08-311-611A-16

```
; Sequence 16, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: "BPI.15"
; US-08-311-611A-16

Query Match 100.0%; Score 55; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKKRELK 10
Db 6 KWKAKKRELK 15

RESULT 13
US-08-311-611A-17
; Sequence 17, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: "BPI.16"
; US-08-311-611A-17

Query Match 100.0%; Score 55; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKKRELK 10
Db 6 KWKAKKRELK 15

RESULT 14
US-08-311-611A-18
; Sequence 18, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311.611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273.401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/125.651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "BPI.17"  
US-08-311-611A-18

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10  
Db 6 KWKAKRFLK 15

RESULT 15  
US-08-311-611A-19  
Sequence 19, Application US/08311611A  
Patent No. 5523288  
GENERAL INFORMATION:  
APPLICANT: Cohen, Jonathan  
APPLICANT: Kung, Ada H.C.  
APPLICANT: Lambert, Jr., Lewis H.  
TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial  
TITLE OF INVENTION: Infection by Administration of  
TITLE OF INVENTION: Bactericidal/Permeability-Increasing  
TITLE OF INVENTION:  
NUMBER OF SEQUENCES: 227  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311.611A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/273.401  
FILING DATE: 11-JUL-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/125.651  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharp, Jeffrey S.  
REGISTRATION NUMBER: 31,879  
REFERENCE/DOCKET NUMBER: 32251  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "BPI.18"  
US-08-311-611A-19

Query Match 100.0%; Score 55; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAKRFLK 10  
Db 6 KWKAKRFLK 15

Search completed: September 7, 1999, 23:23:08  
Job time: 912 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5

Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:18:22 ; Search time 40.95 Seconds  
(without alignments)  
5.784 Million cell updates/sec

Title: US-09-124-280A-44

Perfect score: 55

Sequence: 1 KWAQKRELFK 10

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A\_Geneseq\_36:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	487	1 P94636	Human bactericidal
2	55	100.0	368	1 R21840	Sequence of p337,
3	55	100.0	396	1 R21843	Sequence of CHO-BP
4	55	100.0	456	1 R21844	Sequence of BPI (D
5	55	100.0	487	1 R24777	Sequence of human
6	55	100.0	368	1 R33776	BPI from p337. Co
7	55	100.0	243	1 R33777	BPI from p212. Co
8	55	100.0	487	1 R44420	PING4512 encoded f
9	55	100.0	487	1 R43886	Human bactericidal
10	55	100.0	35	1 R63687	BPI derived peptid
11	55	100.0	487	1 R57901	Bactericide BPI. B
12	55	100.0	487	1 R62370	Recombinant bacter
13	55	100.0	20	1 R63731	BPI derived peptid
14	55	100.0	15	1 R63727	BPI derived peptid
15	55	100.0	20	1 R63730	BPI derived peptid
16	55	100.0	24	1 R63729	BPI derived peptid
17	55	100.0	21	1 R63728	BPI derived peptid
18	55	100.0	26	1 R63735	BPI derived peptid
19	55	100.0	24	1 R62495	BPI derived peptid
20	55	100.0	29	1 R62496	BPI derived peptid
21	55	100.0	24	1 R62497	BPI derived peptid
22	55	100.0	15	1 R63715	BPI derived peptid
23	55	100.0	17	1 R63720	BPI derived peptid
24	55	100.0	10	1 R63689	BPI derived peptid
25	55	100.0	16	1 R63690	BPI derived peptid
26	55	100.0	17	1 R63691	BPI derived peptid
27	55	100.0	27	1 R63692	BPI derived peptid
28	55	100.0	15	1 R63693	BPI derived peptid
29	55	100.0	15	1 R63694	BPI derived peptid
30	55	100.0	15	1 R63695	BPI derived peptid
31	55	100.0	15	1 R63696	BPI derived peptid
32	55	100.0	15	1 R63697	BPI derived peptid
33	55	100.0	487	1 R62344	Recombinant 23 kD
34	55	100.0	15	1 R63688	BPI derived peptid
35	55	100.0	487	1 R67998	Holoprotein RBP150
36	55	100.0	487	1 R71290	Human bactericidal
37	55	100.0	368	1 R71291	Human bactericidal
38	55	100.0	243	1 R71292	Human bactericidal
39	55	100.0	10	1 R71798	Peptide neutralisi
40	55	100.0	199	1 R68925	BPI(1-159)/lipopol
41	55	100.0	197	1 R68929	Lipopolysaccharide
42	55	100.0	199	1 R68931	LBP(1-87)/BPI(88-1
43	55	100.0	198	1 R68933	BPI(1-147)/LBP(147

## ALIGNMENTS

RESULT 1

P94636 standard; protein; 487 AA.

AC P94636;

DT 29-JUN-1990 (first entry)

DE Human bactericidal/permeability-increasing (BPI) holoprotein

KW Polypeptide fragment; bacteraemia; sepsis; gram-negative bacteria;

KW bactericidal/permeability-increasing (BPI) holoprotein;

KW mammalian myeloid blood cells; polymorphonuclear leukocytes (PMN);

KW human promyelocytic leukaemia cells (HL-60).

OS Homo Sapiens.

FH Key Location/Qualifiers

FT peptide 1..31

FT protein 32..487

FT region 153..155

FT /note="potential glycosylation site"

FT region 379..381

FT /note="potential glycosylation site"

FN W08901486-A.

PD 23-FEB-1989.

PF 09-AUG-1988; U02700.

PR 05-AUG-1988; US-228035, US-084335.

PA (UNY) New York Univ.

PI Elsbach P, Weiss J;

DR WPI; 89-068849/09.

DR N-PSDB; N92766

PT Bacterial-permeability increasing holoprotein fragments -

PT used for treatment of diseases caused by gram negative bacteria

PT such as bacteraemia or sepsis.

PS Disclosure; ; 41pp; English.

CC The purified, isolated cDNA which encodes this sequence was isolated

CC from human promyelocytic leukaemia cells (HL-60) and is also claimed.

CC Biologically active BPI protein fragments can be used for the

CC treatment of mammals suffering from diseases caused by gram-negative

CC bacteria such as bacteraemia or sepsis. They have good selectivity and

CC lack of cytotoxicity towards cells other than gram-negative bacteria.

CC The fragments may be used in a prepn. further comprising rifampicin,

CC penicillin or cephalosporin. The BPI holoproteins used in prodn. of

CC the fragments may be obt'd from mammalian myeloid blood cells, eg

CC polymorphonuclear leukocytes (PMN), or by recombinant DNA techniques.

CC BPI fragments are prep'd. by incubating the holoprotein under protein

CC cleaving conditions or using recombinant DNA techniques.

SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.061;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWAQKRELFK 10

Db 121 KWAQKRELFK 130

RESULT 2

R21840 standard; protein; 368 AA.

AC R21840;

DT 22-JUL-1992 (first entry)

DE Sequence of p337, a biologically active fragment of the human

DE 57kd protein which binds to the outer membrane of susceptible gram

DE negative bacteria (BPI).

DE Endotoxin binding protein; diagnosis; therapy; endotoxemia;

KW Lipopolysaccharide.

OS Homo sapiens.

FH Key Location/Qualifiers

FT region 1..31  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI; 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endotoxaemia and endo-toxic shock  
 PS Disclosure: Fig 13; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 55; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRELK 10  
 |||||  
 DB 121 KWKAKQKRELK 130

RESULT 3  
 R21843  
 ID R21843 standard; Protein; 396 AA.  
 AC R21843;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of CHO-BPI a biologically active variant of BPI, a human  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI; 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endotoxaemia and endo-toxic shock  
 PS Claim 5; Fig 24; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 396 AA;

Query Match 100.0%; Score 55; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRELK 10  
 |||||  
 DB 90 KWKAKQKRELK 99

RESULT 4  
 R21844

ID R21844 standard; Protein; 456 AA.  
 AC R21844;  
 DT 22-JUL-1992 (first entry)  
 DE Sequence of BPI (DP linkage), a biologically active variant of BPI, a  
 DE 57kd protein which binds to the outer membrane of susceptible gram  
 DE negative bacteria (BPI).  
 KW Endotoxin binding protein; diagnosis; therapy; endotoxaemia;  
 KW endotoxic shock.  
 OS Homo sapiens.  
 PN WO9203535-A.  
 PD 05-MAR-1992.  
 PF 13-AUG-1991; U05758.  
 PR 13-AUG-1990; US-567016.  
 PR 05-APR-1991; US-681551.  
 PA (INCY-) INCYTE PHARM INC.  
 PI Marra MN, Scott RW;  
 DR WPI; 92-096879/12.  
 PT New recombinant BPI protein - specifically binds to and inhibits  
 PT endotoxin lethality, for prevention, treatment and diagnosis of  
 PT endo-toxaemia and endo-toxic shock  
 PS Claim 52; Fig 25; 108pp; English.  
 CC The BPI binds endotoxin in the presence of serum and plasma and,  
 CC unlike other known endotoxin binding proteins such as LBP  
 CC (lipopolysaccharide binding protein), BPI inhibits the  
 CC immunostimulatory and toxic activities of endotoxin both in vitro  
 CC and in vivo. Thus BPI can be used in the diagnosis and therapeutic  
 CC and prophylactic treatment of endotoxin-related disorders including  
 CC endotoxaemia and endotoxic shock.  
 SQ Sequence 456 AA;

Query Match 100.0%; Score 55; DB 1; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 0.058;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRELK 10  
 |||||  
 DB 90 KWKAKQKRELK 99

RESULT 5  
 R24777  
 ID R24777 standard; Protein; 487 AA.  
 AC R24777;  
 DT 02-JAN-1992 (first entry)  
 DE Sequence of human bactericidal/permeability-increasing protein  
 DE (hBPI) derived from human promyelocytic leukemia cells (HL-60).  
 KW Bactericidal; endotoxin inhibitor; anti-microbial.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..31  
 FT /label= signal  
 FT modified\_site 153..155  
 FT /note= "potential N-linked glycosylation site"  
 FT modified\_site 379..381  
 FT /note= "see above"  
 PN WO9209621-A.  
 PD 11-JUN-1992.  
 PF 03-DEC-1991; U09033.  
 PR 03-DEC-1990; US-621473.  
 PR 26-AUG-1991; US-754204.  
 PA (UINY ) UNIV NEW YORK STATE.  
 PI Elsbach P, Weiss J;  
 DR WPI; 92-217016/26.  
 DR N-PSDS; Q25593.  
 PT Antibacterial fragments of bactericidal-permeability-increasing  
 PT protein - for treating Gram negative bacterial infections esp. in  
 PT immuno suppressed patients, with endotoxin activity  
 PS Claim 16; Fig 5; 63pp; English.

CC The hBPI holoprotein was isolated from human promyelocytic leukemia  
 CC cells (HL-60) and BPI cDNA was prepd. The N-terminal fragments of  
 CC BPI exhibit all the anti-bacterial and membrane permeability-  
 CC increasing functions contained in the complete molecule (as well as  
 CC the LPS-neutralising function) but are substantially smaller. In  
 CC addition, the N-terminal fragment, as well as the C-terminal  
 CC fragment(s) are capable of inhibiting the prodn. and/or release of  
 CC tumour necrosis factor (TNF) triggered by LPS in whole blood. The  
 CC inventors claim the DNA sequence comprising bps 721-1491 of Q25593.  
 CC They also claim the following fragments of MATURE hBPI: 1-99 (I);  
 CC 1-203 (II); 200-456 (III); 204-456 (IV). NB: The above fragments  
 CC are labeled with respect to the first AA of the mature sequence,  
 CC which corresp. to residue 32 of R24777.  
 SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
 |||||

Db 121 KWKAQKRELK 130

## RESULT 6

R33776  
 ID R33776 standard; protein; 368 AA.

AC R33776;

DE 22-JUL-1993 (first entry)

KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxemia; adult respiratory distress syndrome.

OS Homo sapiens.

FH Key Location/Qualifiers

FT peptide 1..32

FT region /note= "signal peptide"

FT 32..368

FT /note= "mature protein"

PN W09305797-A.

PD 01-APR-1993.

PF 28-SEP-1992; U08234.

PR 27-SEP-1991; US-766566.

PA (INCY-) INCYTE PHARM INC.

PI Marra MN, Scott RW, Snable JL, Wilde CG;

DR WPI; 93-117241/14.

PT Compsns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an  
 PT active form

PS Disclosure: Fig 6; 53pp; English.

CC The human bactericidal/permeability increasing protein may be used  
 CC in a compn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related -shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33777. R33729-31.  
 SQ Sequence 368 AA;

Query Match 100.0%; Score 55; DB 1; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 0.047;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
 |||||

Db 121 KWKAQKRELK 130

## RESULT 7

R33777  
 ID R33777 standard; protein; 243 AA.

AC R33777;

DE 22-JUL-1993 (first entry)

KW BPI from p212.

KW Bacterial/permeability increasing protein; lipopolysaccharide; LPS;  
 KW inhibitor; endotoxin-related shock; disseminated intravascular  
 KW coagulation; anaemia; thrombocytopenia; ARDS; renal failure;  
 KW endotoxemia; adult respiratory distress syndrome.

OS Homo sapiens.

FH Key Location/Qualifiers

FT peptide 1..32

FT region /note= "signal peptide"

FT 32..243

FT /note= "mature protein"

PN W09305797-A.

PD 01-APR-1993.

PF 28-SEP-1992; U08234.

PR 27-SEP-1991; US-766566.

PA (INCY-) INCYTE PHARM INC.

PI Marra MN, Scott RW, Snable JL, Wilde CG;

DR WPI; 93-117241/14.

PT Compsns. contg. a bactericidal and permeability increasing  
 PT protein - solubilised in a lipid carrier to maintain stably an  
 PT active form

PS Disclosure: Fig 7; 53pp; English.

CC The human bactericidal/permeability increasing protein may be used  
 CC in a compn. solubilised in a lipid carrier. The BPI binds  
 CC lipopolysaccharides and inhibits the immunostimulatory and toxic  
 CC activities of LPS in vitro and in vivo. The lipid compsn. stably  
 CC maintains the BPI in a sol. active form. The compsn. can be used  
 CC for detecting or quantifying LPS in a sample, coating a surgical  
 CC tool, coating an implantable invasive device, decontaminating a  
 CC fluid contg. LPC and treating endotoxin related -shock,  
 CC -disseminated intravascular coagulation, -anaemia, -thrombocytopenia,  
 CC -adult respiratory stress syndrome or -renal failure, or for  
 CC preventing endotoxaemia.  
 CC See also R33776, R33729-31.  
 SQ Sequence 243 AA;

Query Match 100.0%; Score 55; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KWKAQKRELK 10  
 |||||

Db 121 KWKAQKRELK 130

## RESULT 8

R44420

ID R44420 standard; Protein; 487 AA.

AC R44420;

DE 03-JUN-1994 (first entry)

KW PING4512 encoded fusion protein.

KW Polymerase chain reaction; primer; amplify; PCR; plasmid; PING4512;  
 KW PING4503; bactericidal/permeability-increasing protein; BPI; IGG;  
 KW heavy; chain; PMB27; fusion protein; Gram negative; renal failure;  
 KW bacterial infection; endotoxin related shock; metabolic acidosis;  
 KW disseminated intravascular coagulation; anemia; thrombocytopenia;  
 KW leukopenia; adult respiratory distress syndrome; hypotension; fever;  
 KW constant region; lipopolysaccharide complement binding;  
 KW placental transfer; Fc receptor binding.

OS Synthetic.

FH Key Location/Qualifiers

FT peptide 1..31

FT region /note= "Signal peptide"

FT 32..487

FT protein

FT W09323434-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04754.  
 PR 19-MAY-1992; US-885911.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS, Horwitz A, Theofan G;  
 DR WPI; 93-386485/48.  
 DR N-PSDB; Q52488.  
 PT New fusion proteins for treating bacterial infections - comprising  
 PT a bactericidal-permeability-increasing protein and an immunoglobulin  
 PT heavy chain constant domain  
 PS Claim 20; Page 48-50; 75pp; English.  
 CC This sequence is encoded by the plasmid PING4512. PING4512 was  
 CC constructed using the primer sequence given in Q52480. Plasmid  
 CC PING4503 was cut with AlwNI, the ends made blunt with T4 DNA  
 CC polymerase, and the DNA then cut with SalI. The resultant approx.  
 CC 700bp SalI/blunt fragment contains the 30 bp 5' untranslated region  
 CC and the DNA encoding the signal peptide and first 191 amino acids of  
 CC bactericidal/permeability-increasing protein (BPI). IgG heavy chain  
 CC (HC) sequences, which also included the hinge region of the IgG HC  
 CC were PCR amplified from plasmid pMB27 using these primers. Plasmids  
 CC such as PING4512 may be used to produce fusion proteins which are  
 CC useful for the treatment of Gram negative bacterial infections and  
 CC their sequelae including endotoxin related shock and conditions  
 CC associated with it, such as disseminated intravascular coagulation,  
 CC anemia, thrombocytopenia, leukopenia, adult respiratory distress  
 CC syndrome, renal failure, hypotension, fever and metabolic acidosis.  
 CC proving BPI or a fragment of it, as part of the fusion with an  
 CC immunoglobulin heavy chain constant region provides the potential  
 CC advantages of Fc receptor binding, bivalent binding to lipopoly-  
 CC saccharide complement binding and increasing placental transfer.  
 CC Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10  
 |||||  
 DB 121 KWKAQKRFLK 130

## RESULT 9

ID R43886 standard; Protein; 487 AA.  
 AC R43886;  
 DT 01-JUN-1994 (first entry)  
 DE Human bactericidal/permeability-increasing protein.  
 KW Bactericidal/permeability-increasing protein;  
 KW endotoxin binding protein.  
 OS Homo sapiens.  
 FH KEY Location/Qualifiers  
 FT protein  
 FT 32..487  
 FT /label= human BPI protein  
 FT misc\_difference 1..31  
 FT /label= signal peptide  
 PN W09323540-A.  
 PD 25-NOV-1993.  
 PF 19-MAY-1993; U04752.  
 PR 19-MAY-1992; US-885501.  
 PA (XOMA ) XOMA CORP.  
 PI Grinna LS;  
 DR WPI; 93-386572/48.  
 DR P-PSDB; Q52268.  
 PT Recovering endotoxin binding protein from transformed cell  
 PT cultures - by retention on cation exchange material and elution,  
 PT giving simpler recovery and higher yields, for treating bacterial  
 PT infections  
 PS Disclosure; Page 27-28; 38pp; English.  
 CC Human bactericidal/permeability increasing protein (an endotoxin  
 CC binding protein) is useful for treatment of bacterial infections.

SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.061;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10  
 |||||  
 DB 121 KWKAQKRFLK 130

## RESULT 10

ID R63687 standard; peptide; 35 AA.  
 AC R63687;  
 DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, domain II.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipoarabinomannan; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.  
 PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG;  
 DR WPI; 94-302964/37.  
 PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity  
 PS Claim 6; Page 157; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63687-726 are derived from  
 CC positions 65-99 of BPI. Peptides such as these may be used as  
 CC heparin binding agents, for neutralising the anti-coagulant effect  
 CC of heparin, for inhibiting angiogenesis, eg. associated with ocular  
 CC retinopathy, for inhibiting endothelial cell proliferation, for  
 CC for treating a chronic inflammatory tumour cell proliferation,  
 CC for treating a gram-negative disease state, eg. rheumatoid  
 CC arthritis, and for treating gram-negative bacterial infection. The  
 CC peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipoarabinomannan in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipoarabinomannan or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 35 AA;

Query Match 100.0%; Score 55; DB 1; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 0.0057;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10  
 |||||  
 DB 26 KWKAQKRFLK 35

```

RESULT 11
ID R57901 standard; Protein; 487 AA.
AC R57901;
DT 29-MAR-1995 (first entry)
DE Bactericide BPI.
KW BPI: bactericidal-permeability-increasing protein; bactericide;
OS fusion protein; Gram-negative bacterium; infection.
FH Homo sapiens.
FT Key Location/Qualifiers
FT peptide 1..31
FT /label= Sig_peptide
FT WO9418323-A.
PN 18-AUG-1994.
PF 02-FEB-1994; U01235.
PR 02-FEB-1993; US-013801.
PA (XOMA ) XOMA CORP
PI Baftian M, Burke D, Grinna L, Horwitz A, Theofan G;
DR WPI: 94-279744/34.
DR N-PSDB; Q67270.
PT Bactericidal-permeability-increasing protein analogs with
PT Cys132 or Cys135 replaced - also fusion proteins and C-terminally
PT truncated forms, have increased stability and are useful to treat
PT gram-negative bacterial infection
PS Disclosure; Page 52:54; 77pp; English.
CC Analogs, fusion proteins and C-terminal truncated forms of human BPI
CC are claimed that show improved resistance to dimerization and
CC adduct formation.
SQ Sequence 487 AA;

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10
DB 121 KWKAQKRFLK 130
|||||

```

```

RESULT 12
ID R62370 standard; Protein; 487 AA.
AC R62370;
DT 27-APR-1995 (first entry)
DE Recombinant bactericidal/permeability-increasing protein (rBPI-50).
KW N-terminal; recombinant; bactericidal/permeability protein; rBPI;
KW signal protein; human; rBPI-23; binding assay; heparin-binding; domain;
KW Limulus amoebocyte lysate; LAL; inhibition; bactericidal activity;
KW chronic inflammatory disease; arthritis; psoriasis; Crohn's disease;
KW inflammatory bowel disease; asthma.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..31
FT /note= "Signal peptide"
FT protein 32..487
FT /note= "mature protein"
FT domain 17..55
FT /note= "LAL inhibiting domain"
FT binding_site 21..55
FT /note= "Heparin binding site"
FT binding_site 65..107
FT /note= "Heparin binding site"
FT domain 73..99
FT /note= "LAL inhibiting domain"
FT active_site 85..99
FT /note= "Bactericidally active fragment"
FT domain 137..163
FT /note= "LAL inhibiting domain"
FT binding_site 137..171
FT /note= "Heparin binding site"
PN WO9420128-A.
PD 15-SEP-1994.

Query Match 100.0%; Score 55; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAQKRFLK 10
DB 121 KWKAQKRFLK 130
|||||

```

```

RESULT 13
R63731
ID R63731 standard; peptide; 20 AA.
AC R63731;
DT 03-MAY-1995 (first entry)
DE BPI derived peptide, BPI.47.
KW Human; bactericidal/permeability-increasing protein; BPI; heparin;
KW binding agent; neutralisation; anti-coagulant effect; inhibition;
KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;
KW contraception; malignant; tumour cell; inflammatory disease; cytokine;
KW rheumatoid arthritis; gram-negative bacteria; infection; microbe;
KW liposarabinomannan; circulation; compromised immune response; macrophage; activation; lymphokine; decontaminating; Helicobacter;
KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;
KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.
OS Homo sapiens.
PN WO9420532-A.
PD 15-SEP-1994.
PF 11-MAR-1994; U02465.
PR 12-MAR-1993; US-030644.
PR 15-JUL-1993; US-093202.
PR 14-JAN-1994; US-183222.
PA (XOMA ) XOMA CORP.
PI Little RG;
DR WPI: 94-302964/37.
PT New human bactericidal permeability increasing peptides - derived
PT from the functional domains of BPI and having BPI activities such
PT as bactericidal activity
PS Claim 8; Page 159; 254pp; English.
CC The sequences given in R63682-750, R62087-100 and R62491-500 are
CC peptides derived from human bactericidal/permeability-increasing
CC protein (BPI). The sequences given in R63727-35 are derived from
CC positions 65-99 of BPI and may form dimers or trimers. Peptides
CC such as these may be used as heparin binding agents, for neutralising
CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.
CC associated with ocular retinopathy, for inhibiting endothelial cell
CC proliferation, for contraception, for inhibiting malignant tumour cell
CC proliferation, for treating a chronic inflammatory disease state, eg.

```

CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipoarabinomannan in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipoarabinomannan or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 20 AA:

Query Match 100.0%; Score 55; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRFLLK 10  
 | | | | | | | | | |  
 Db 1 KWKAKQKRFLLK 10

## RESULT 14

ID R63727 standard; peptide: 15 AA.

AC R63727;  
 DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, BPI.9.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipoarabinomannan; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.

PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.  
 PA (XOMA ) XOMA CORP.  
 PI Little RG.

DR WPI; 94-302964/37.  
 PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity

PS Claim 8; Page 158; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63727-35 are derived from  
 CC positions 65-99 of BPI and may form dimers or trimers. Peptides  
 CC such as these may be used as heparin binding agents, for neutralising  
 CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.  
 CC associated with ocular retinopathy, for inhibiting endothelial cell  
 CC proliferation, for contraception, for inhibiting malignant tumour cell  
 CC proliferation, for treating a chronic inflammatory disease state, eg.  
 CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipoarabinomannan in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipoarabinomannan or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 15 AA:

Query Match 100.0%; Score 55; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRFLLK 10  
 | | | | | | | | | |  
 Db 6 KWKAKQKRFLLK 15

## RESULT 15

ID R63730 standard; peptide: 20 AA.

AC R63730; 1995 (first entry)

DT 03-MAY-1995 (first entry)  
 DE BPI derived peptide, BPI.46.  
 KW Human; bactericidal/permeability-increasing protein; BPI; heparin;  
 KW binding agent; neutralisation; anti-coagulant effect; inhibition;  
 KW angiogenesis; ocular retinopathy; endothelial cell; proliferation;  
 KW contraception; malignant; tumour cell; inflammatory disease; T-cell;  
 KW rheumatoid arthritis; gram-negative bacteria; infection; cytokine;  
 KW lipoarabinomannan; circulation; compromised immune response; microbe;  
 KW macrophage; activation; lymphokine; decontaminating; Helicobacter;  
 KW gastritis; peptic ulcer; gastric ulcer; duodenal ulcer; antibiotic;  
 KW gentamicin; polymyxin B; cefamandole nafate; LBP protein.  
 OS Homo sapiens.  
 PN W09420532-A.  
 PD 15-SEP-1994.

PF 11-MAR-1994; U02465.  
 PR 12-MAR-1993; US-030644.  
 PR 15-JUL-1993; US-093202.  
 PR 14-JAN-1994; US-183222.

PA (XOMA ) XOMA CORP.

PI Little RG.

DR WPI; 94-302964/37.

PT New human bactericidal permeability increasing peptides - derived  
 PT from the functional domains of BPI and having BPI activities such  
 PT as bactericidal activity

PS Claim 8; Page 159; 254pp; English.  
 CC The sequences given in R63682-750, R62087-100 and R62491-500 are  
 CC peptides derived from human bactericidal/permeability-increasing  
 CC protein (BPI). The sequences given in R63727-35 are derived from  
 CC positions 65-99 of BPI and may form dimers or trimers. Peptides  
 CC such as these may be used as heparin binding agents, for neutralising  
 CC the anti-coagulant effect of heparin, for inhibiting angiogenesis, eg.  
 CC associated with ocular retinopathy, for inhibiting endothelial cell  
 CC proliferation, for contraception, for inhibiting malignant tumour cell  
 CC proliferation, for treating a chronic inflammatory disease state, eg.  
 CC rheumatoid arthritis, and for treating gram-negative bacterial infection.  
 CC The peptides may also be used for treating a subject suffering from the  
 CC adverse effects of the presence of lipoarabinomannan in the circulation,  
 CC eg. a compromised immune response to microbes or tumour cells due to  
 CC inhibition of macrophage activation by T-cell lymphokines or increased  
 CC production of a cytokine, for decontaminating a fluid containing  
 CC lipoarabinomannan or for treating a disease associated with Helicobacter  
 CC infection, eg. gastritis, peptic ulcer, gastric ulcer or duodenal ulcer.  
 CC The peptides can be used with an antibiotic eg. gentamicin, polymyxin B  
 CC or cefamandole nafate or LBP protein products. The peptides are pref.  
 CC prepared by solid phase synthesis.  
 CC Sequence 20 AA:

Query Match 100.0%; Score 55; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.0035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KWKAKQKRFLLK 10  
 | | | | | | | | | |  
 Db 11 KWKAKQKRFLLK 20



Wed Sep 8 09:45:21 1999

us-09-124-280a-44.rag

Page 7

Search completed: September 7, 1999, 23:18:23  
Job time: 1679 sec

**THIS PAGE BLANK (USPTO)**

P09092 mus musculu  
P15861 brachydanio

Run On: September 8, 1999, 00:43:02 ; Search time 21.13 Seconds  
(without alignments)  
14.716 Million cell updates/sec

Perfect score: 64

Scoring table: PROGRES

Searched: 77977 seqs, 28268293 residues

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query			DB	ID	Description
		Match	Length	Q			
1	64	100.0	102	1	ALPS_TACTR	P07087 tachypneus	
2	61	95.3	101	1	ALPS_LIMPO	P07086 limulus pol	
3	45	70.3	668	1	F13B_MOUSE	Q07968 mus musculus	
4	42	65.6	661	1	F13B_HUMAN	P05160 homo sapien	
5	41	64.1	563	1	HNMI_YEAST	P19807 saccharomyc	
6	40	62.5	2222	1	DPOE_YEAST	P21951 saccharomyc	
7	40	62.5	277	1	HXDB_NOTVI	P31263 notophthalm	
8	39	60.9	2029	1	LAR_DROME	P16621 drosophila	
9	39	60.9	106	1	YQ95_YEAST	P32339 saccharomyc	
10	38	59.4	323	1	CD47_HUMAN	Q08722 homo sapien	
11	38	59.4	215	1	RL10_EUGGR	Q39724 euglena gra	
12	38	59.4	527	1	UDAL_RAT	P36510 rattus norv	
13	37	57.8	280	1	HXDB_CHICK	P24342 gallus gall	
14	37	57.8	347	1	NUZM_DASNO	O21326 dasyopus nov	
15	37	57.8	273	1	PSAB_YERPE	P31523 yersinia pe	
16	36	56.2	1255	1	CD45_RAT	P04157 rattus norv	
17	36	56.2	1087	1	PGDS_XENLA	P26619 xenopus lae	
18	36	56.2	766	1	PRTP_HSVEB	P28973 equine herp	
19	36	56.2	337	1	RT07_ACACA	P46756 acanthamoeb	
20	35	54.7	323	1	HXDB_MOUSE	P23813 mus musculus	
21	35	54.7	859	1	OSHI_YEAST	P35845 saccharomyc	
22	35	54.7	355	1	PYRC_SULSO	Q08357 sulfilobus	
23	35	54.7	329	1	RIR2_BACSU	P50621 bacillus su	
24	35	54.7	737	1	UGS1_HUMAN	P13807 homo sapien	
25	35	54.7	737	1	UGS3_MOUSE	P54859 mus musculus	
26	35	54.7	232	1	VIF_SIVAI	P27974 simian immu	
27	35	54.7	161	1	Y447_METJA	Q57889 methanococc	
28	35	54.7	1756	1	YCF1_PINTH	P41647 pinus thunb	
29	35	54.7	389	1	YG2X_YEAST	P53266 saccharomyc	
30	35	54.7	954	1	YJ59_YEAST	P47134 saccharomyc	
31	34.5	53.9	813	1	DD13_HUMAN	O43131 homo sapien	
32	34.5	53.9	758	1	DD13_MOUSE	O35286 mus musculus	
33	34	53.1	401	1	ACKA_HAEIN	P44406 haemophilus	
34	34	53.1	232	1	CSC1_ECOLI	P53518 escherichia	
35	34	53.1	232	1	CSC2_ECOLI	P53519 escherichia	
36	34	53.1	2257	1	DPOE_HUMAN	Q07864 homo sapien	
37	34	53.1	220	1	HB7A_XENLA	Q91771 xenopus lae	
38	34	53.1	220	1	HB7B_XENLA	P04476 xenopus lae	
39	34	53.1	113	1	HEMT_PHAGO	P02244 phascocolpsi	
40	34	53.1	113	1	HEMT_STPCU	P22766 siphonoma	
41	34	53.1	113	1	HEMT_THEZO	P02425 themiste zo	
42	34	53.1	66	1	HXA6_DROME	P02833 drosophila	
43	34	53.1	378	1	HXA6_HUMAN	P31267 homo sapien	

CC LPS-MEDIATED ACTIVATION OF THE HEMOLYMPH COAGULATION. IT HAS A  
 CC STRONG ANTIBACTERIAL EFFECT ESPECIALLY ON THE GROWTH OF GRAM-  
 CC NEGATIVE BACTERIA.  
 CC PIR: A27819; A27819.

KW ANTIBIOTIC. 31 52

FT DISULFID 13 13 K -> N.

FT VARIANT 13 13  
 SQ SEQUENCE 101 AA; 11801 MW; 32EB8EF CRC32;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0004;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11

DB 40 RLKWKYKGF 50

RESULT 3

ID FL13B\_MOUSE STANDARD; PRT; 668 AA.

AC Q07968;

DT 01-FEB-1995 (REL. 31, CREATED)

DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE COAGULATION FACTOR XIII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-

DE GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B

DE CHAIN)

GN FL13B OR CF13B.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

CC [1]

RP SEQUENCE FROM N.A.

RC STRAIN-B10.D2/OSN; TISSUE=LIVER;

RX MEDLINE; 93224141.

RA NONAKA M., MATSUDA Y., SHIROISHI T., MORIWAKI K., NONAKA M.,

RA NATSUUME-SAKAI S.;

RT "Molecular cloning of the b subunit of mouse coagulation factor XIII

RT and assignment of the gene to chromosome 1: close evolutionary

RT relationship to complement factor H.;"

RL GENOMICS 15:535-542(1993).

CC -!- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,

CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE

CC OF TRANSGLUTAMINASE FORMATION BY THROMBIN.

CC -!- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER AND KIDNEY.

CC -!- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; D10071; G303652; "

DR PIR: A46013; A46013.

DR MGD; MGI:88379; CF13B.

DR PFAM; PF00084; sushi; 8.

DR HSSP; P10998; 1VVC.

KW TRANSFERASE; PLASMA; BLOOD COAGULATION; REPEAT; GLYCOPROTEIN; SIGNAL;

SUSHI.

FT SIGNAL 1 20

FT CHAIN 21 668 BY SIMILARITY.

FT DOMAIN 24 647 FACTOR XIII, B CHAIN.

FT REPEAT 24 88 10 X SUSHI (SCR) REPEATS.

FT REPEAT 24 88 SUSHI 1.

FT REPEAT 90 147 SUSHI 2.

FT REPEAT 152 209 SUSHI 3.

FT REPEAT 212 268 SUSHI 4.

FT REPEAT 273 328 SUSHI 5.

FT REPEAT 335 390

FT REPEAT 395 451 SUSHI 6.

FT REPEAT 453 516 SUSHI 7.

FT REPEAT 523 579 SUSHI 8.

FT REPEAT 581 647 SUSHI 9.

FT DISULFID 25 76 SUSHI 10.

FT DISULFID 59 87 BY SIMILARITY.

FT DISULFID 91 135 BY SIMILARITY.

FT DISULFID 118 146 BY SIMILARITY.

FT DISULFID 153 197 BY SIMILARITY.

FT DISULFID 180 208 BY SIMILARITY.

FT DISULFID 213 255 BY SIMILARITY.

FT DISULFID 241 267 BY SIMILARITY.

FT DISULFID 274 316 BY SIMILARITY.

FT DISULFID 302 327 BY SIMILARITY.

FT DISULFID 336 378 BY SIMILARITY.

FT DISULFID 364 389 BY SIMILARITY.

FT DISULFID 396 439 BY SIMILARITY.

FT DISULFID 425 450 BY SIMILARITY.

FT DISULFID 454 505 BY SIMILARITY.

FT DISULFID 486 515 BY SIMILARITY.

FT DISULFID 524 567 BY SIMILARITY.

FT DISULFID 553 578 BY SIMILARITY.

FT DISULFID 582 636 BY SIMILARITY.

FT DISULFID 616 646 BY SIMILARITY.

FT CARBOHYD 162 162 POTENTIAL.

FT CARBOHYD 545 545 POTENTIAL.

SQ SEQUENCE 668 AA; 76078 MW; 45234BA0 CRC32;

Query Match

Best Local Similarity 70.3%; Score 45; DB 1; Length 668;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 10

DB 467 QLKWKYKGF 475

RESULT 4

ID FL13B\_HUMAN

AC P05160; STANDARD; PRT; 661 AA.

DT 13-AUG-1987 (REL. 05, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE COAGULATION FACTOR XIII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-

DE GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B

DE CHAIN).

GN FL13B.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

CC [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91105054.

RA BOTTENUS R.E., ICHINOSE A., DAVIE E.W.;

RT "Nucleotide sequence of the gene for the b subunit of human factor

RT XIII.;"

RL BIOCHEMISTRY 29:11195-11209(1990).

RP SEQUENCE OF 2-661 FROM N.A.

RX MEDLINE; 87026535.

RA ICHINOSE A., MCMULLEN B.A., FUJIKAWA K., DAVIE E.W.;

RT "Amino acid sequence of the b subunit of human factor XIII, a protein

RT composed of ten repetitive segments.;"

RL BIOCHEMISTRY 25:4633-4638(1986).

CC [3]

RP REVISIONS.

RA ICHINOSE A.;

RL SUBMITTED (FEB-1987) TO EMBL/GENBANK/DBJ DATA BANKS.

CC [4]

RP SEQUENCE OF 1-20 FROM N.A.

RC TISSUE-LIVER;  
 RX MEDLINE; 90251457.  
 RA GRUNDMANN U., NERLICH C., REIN T., ZETTLMEISSL G.;  
 RT "Complete cDNA sequence encoding the B subunit of human factor XIII.";  
 RL NUCLEIC ACIDS RES. 18:2817-2817(1990).  
 RN [5]  
 RP VARIANT PHE-450  
 RX MEDLINE; 93313189.  
 RA HASHIGUCHI T., SAITO M., MORISHITA E., MATSUDA T., ICHINOSE A.;  
 RT "Two genetic defects in a patient with complete deficiency of the b-  
 RL subunit for coagulation factor XIII.";  
 RL BLOOD 82:145-150(1993).  
 CC -1- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,  
 CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE  
 CC OF TRANSGLUTAMINASE FORMATION BY THROMBIN.  
 CC -1- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.  
 CC -1- DISEASE: A DEFICIENCY IN FAL3 CAN RESULT IN A LIFE-LONG BLEEDING  
 CC TENDENCY, DEFECTIVE WOUND HEALING, AND HABITUAL ABORTION.  
 CC -1- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M64554; G179417; ALT\_SEQ.  
 DR EMBL; M14057; G182839;  
 DR EMBL; X51823; G31505;  
 DR PIR; A23830; A23830.  
 DR PIR; A36397; A36397.  
 DR PIR; S09980; S09980.  
 DR MIM; 134580;  
 DR PFAM; PF00084; sushi; 8.  
 DR HSP; P08603; IPI.  
 KW TRANSFERASE; PLASMA; BLOOD COAGULATION; REPEAT; GLYCOPROTEIN; SIGNAL;  
 SUSHI; DISEASE MUTATION.  
 FT SIGNAL 1 20  
 FT CHAIN 21 661  
 FT DOMAIN 24 647  
 FT REPEAT 24 88  
 FT REPEAT 90 147  
 FT REPEAT 152 209  
 FT REPEAT 212 268  
 FT REPEAT 273 328  
 FT REPEAT 335 390  
 FT REPEAT 395 451  
 FT REPEAT 453 516  
 FT REPEAT 523 579  
 FT REPEAT 581 647  
 FT REPEAT 25 76  
 FT DISULFID 59 87  
 FT DISULFID 91 135  
 FT DISULFID 118 146  
 FT DISULFID 153 197  
 FT DISULFID 180 208  
 FT DISULFID 213 255  
 FT DISULFID 241 267  
 FT DISULFID 274 316  
 FT DISULFID 302 327  
 FT DISULFID 336 378  
 FT DISULFID 364 389  
 FT DISULFID 396 439  
 FT DISULFID 425 450  
 FT DISULFID 454 505  
 FT DISULFID 486 515  
 FT DISULFID 524 567  
 FT DISULFID 553 578  
 FT DISULFID 582 636  
 FT DISULFID 616 646  
 FT CARBOHYD 162

FT CARBOHYD 545 545 POTENTIAL.  
 FT SITE 619 CELL ATTACHMENT SITE.  
 FT VARIANT 450 450 C -> F (IN F13B DEFICIENCY).  
 SQ SEQUENCE 661 AA; 75491 MW; FDD245E9 CRC32;  
 Query Match 55.6%; Score 42; DB 1; Length 661;  
 Best Local Similarity 75.0%; Pred. No. 5.3;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 LKWKYK 10  
 Db 468 MKWYK 475  
 :||||:|  
 RESULT 5  
 HNWL\_YEAST STANDARD; PRT; 563 AA.  
 ID HNWL\_YEAST  
 AC P19807;  
 DT 01-FEB-1991 (REL. 17, CREATED)  
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE CHOLINE TRANSPORT PROTEIN.  
 GN HNWL OR CTRL OR CTR OR IGL077C.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-X2180-1B;  
 RX MEDLINE; 90368823.  
 RA NIKAWA J.-I., HOSAKA K., TSUKAGOSHI Y., YAMASHITA S.;  
 RT "Primary structure of the yeast choline transporter gene and regulation  
 of its expression".  
 RL J. BIOL. CHEM. 265:15996-16003(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 97435481.  
 RA RIEGER M., BRUECKNER M., SCHAEFER M., MUELLER-AUER S.;  
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
 chromosome VII.";  
 RL YEAST 13:1077-1090(1997).  
 CC -1- FUNCTION: SOLE CHOLINE TRANSPORTER IN YEAST.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- INHIBITION OF ACTIVITY BY INTRACELLULAR CHOLINE.  
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J05603; G171330;  
 DR EMBL; Z72599; E243965;  
 DR PIR; S11175; S11175.  
 DR SGD; L0000795; HNWL.  
 DR PROSITE; PS00218; AMINO\_ACID\_PERMEASE; 1.  
 DR PFAM; PF00324; aa\_permeases; 1.  
 KW TRANSPORT; AMINO-ACID TRANSPORT; TRANSMEMBRANE.  
 FT TRANSMEM 58 78 POTENTIAL.  
 FT TRANSMEM 88 108 POTENTIAL.  
 FT TRANSMEM 183 203 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 256 276 POTENTIAL.  
 FT TRANSMEM 294 314 POTENTIAL.  
 FT TRANSMEM 343 363 POTENTIAL.  
 FT TRANSMEM 399 417 POTENTIAL.  
 FT TRANSMEM 427 445 POTENTIAL.  
 FT TRANSMEM 466 486 POTENTIAL.

```

FT TRANSMEM 492 512 POTENTIAL.
FT CARBOHYD 7 7 POTENTIAL.
FT CARBOHYD 20 20 POTENTIAL.
FT CARBOHYD 248 248 POTENTIAL.
FT CARBOHYD 341 341 POTENTIAL.
SQ SEQUENCE 563 AA; 62056 MW; 2ED2DAA7 CRC32;

Query Match 64.1%; Score 41; DB 1; Length 563;
Best Local Similarity 87.5%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LKWKYK 10
DB 508 LKWKYK 515

RESULT 6
DPOE_YEAST STANDARD; PRT; 2222 AA.
AC P21951;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE EPSILON, CATALYTIC SUBUNIT A (EC 2.7.7.7) (DNA
DE POLYMERASE II SUBUNIT A).
GN POL2 OR DUN2 OR YNL262W OR N0825.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1214-1221.
RX MEDLINE: 90381771.
RA MORRISON A., ARAKI H., CLARK A.B., HAMATAKE R.K., SUGINO A.;
RT "A third essential DNA polymerase in S. cerevisiae.";
RL CELL 62:1143-1151(1990).
[2]
SEQUENCE OF 1-2221 FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE: 96310631.
RA SEN-GUPTA M., LYCK R., FLEIG U., NIEDENTHAL R.K., HEGEMANN J.H.;
RT "The sequence of a 24,152 bp segment from the left arm of chromosome
RT XIV from Saccharomyces cerevisiae between the BN11 and the POL2
RT genes.";
RL YEAST 12:505-514(1996).
[3]
TEMPERATURE SENSITIVE MUTANTS.
RX MEDLINE: 92164663.
RA ARAKI H., ROPP P.A., JOHNSON A.L., JOHNSTON L.H., MORRISON A.,
RA SUGINO A.;
RT "DNA polymerase II, the probable homolog of mammalian DNA polymerase
RT epsilon, replicates chromosomal DNA in the yeast Saccharomyces
RT cerevisiae.";
RL EMBO J. 11:733-740(1992).
CC -!- FUNCTION: DNA POLYMERASE II PARTICIPATES IN CHROMOSOMAL DNA
CC REPLICATION.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBUNIT: CONSISTS OF FIVE SUBUNITS (200 KD, 80 KD, 34 KD, 30 KD,
CC AND 29 KD).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE
CC N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY
CC FOR COMPLEXING SUBUNITS B AND C.
CC -!- IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES: ALPHA, BETA, GAMMA,
CC DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR DIFFERENT REACTIONS
CC OF DNA SYNTHESIS.
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. HIGH
CC SIMILARITY WITH MAMMALIAN DNA POLYMERASE EPSILON.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: X57674; -; NOT_ANNOTATED_CDS.
DR PIR: A43783; A43783.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
DR PFAM: PF00046; homeobox; 1.
-----
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: M60416; G171409; -
DR EMBL: X92494; G1045247; -
DR EMBL: 271538; E239648; -
DR PIR: A36028; A36028.
DR SGD: L0001461; POL2.
DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
DR PFAM: PF00136; DNA_POL_B; 3.
KW TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION;
KW DNA-BINDING; ZINC-FINGER; NUCLEAR PROTEIN.
FT ZN_FING 2108 2181
FT SIMILAR 349 388
FT SIMILAR 604 668
FT SIMILAR 768 786
FT SIMILAR 817 858
FT SIMILAR 871 879
FT SIMILAR 976 992
FT VARIANT 644 644 M -> I (IN POL2-9 TS MUTANT).
FT VARIANT 710 710 P -> S (IN POL2-18 TS MUTANT).
SQ SEQUENCE 2222 AA; 255669 MW; 233EFA0E CRC32;

Query Match 62.5%; Score 40; DB 1; Length 2222;
Best Local Similarity 45.5%; Pred. No. 40;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLWKYK 11
DB 679 KRLWKYK 689

RESULT 7
HXDB_NOTVI STANDARD; PRT; 277 AA.
AC P31263;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE HOMEBOX PROTEIN HOX-D11 (NVHox-2).
GN HOXD11 OR NVHox2.
OS NOTOPHTHALMUS VIRIDESCENS (EASTERN NEWT) (TRITURUS VIRIDESCENS).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA;
OC CAUDATA; SALAMANDROIDEA; SALAMANDRIDAE; NOTOPHTHALMUS.
[1]
SEQUENCE FROM N.A.
RX MEDLINE: 91372146.
RA BROWN R., BROCKES J.P.;
RT "Identification and expression of a regeneration-specific homeobox
RT gene in the newt limb blastema.";
RL DEVELOPMENT 111:489-496(1991).
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE ABD-B FAMILY OF HOMEBOX PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: X57674; -; NOT_ANNOTATED_CDS.
DR PIR: A43783; A43783.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
DR PFAM: PF00046; homeobox; 1.

```

DR HSSP: P02835; 1ETZ.  
DR TRANSAC; T01761; -  
KW HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN;  
KW TRANSCRIPTION REGULATION.  
FT DNA\_BIND 205 264 HOMEBOX.  
SQ SEQUENCE 277 AA; 31782 MW; A3CBF4BD CRC32;

Query Match 62.5%; Score 40; DB 1; Length 277;

Best Local Similarity 45.5%; Pred. No. 4.8;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KRLWKYKGF 11  
: | | | | | | |  
Db 68 RRSKWYRGYS 78

RESULT 8  
LAR\_DROME STANDARD; PRT; 2029 AA.  
ID LAR\_DROME  
AC P15621;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE PROTEIN-TYROSINE PHOSPHATASE DLAR PRECURSOR (EC 3.1.3.48) (PROTEIN-TYROSINE-PHOSPHATASE PHOSPHOHYDROLASE).  
GN LAR.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90046860.  
RA STREULI M., KRUEGER N.X., TSAI A.Y.M., SAITO H.;  
RT "A family of receptor-linked protein tyrosine phosphatases in humans and Drosophila.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:8698-8702(1989).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CANTON-S;  
RX MEDLINE; 96178473.  
RA KRUEGER N.X., VAN VACTOR D., WAN H.I., GELBART W.M., GOODMAN C.S., SAITO H.;  
RT "The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila.";  
RL CELL 84:611-622(1996).

CC -1- FUNCTION: IT IS POSSIBLE THAT DLAR IS A CELL ADHESION RECEPTOR. IT POSSESSES AN INTRINSIC PROTEIN TYROSINE PHOSPHATASE ACTIVITY (PTPASE). IT CONTROLS MOTOR AXON GUIDANCE.  
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN A SUBSET OF AXONS AND PIONEER NEURONS IN THE EMBRYO.  
CC -1- SIMILARITY: EXTRACELLULAR REGION TYPIC OF A CAM FAMILY (3 IG-LIKE DOMAINS FOLLOWED BY 9 FIBRONECTIN TYPE III-LIKE DOMAINS), AND A CYTOPLASMIC REGION COMPOSED OF TWO PTPASE DOMAINS.  
CC -1- SIMILARITY: TO DPTP AND TO LAR.

-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M27700; G157812; -  
EMBL; U36857; G1209648; -  
DR EMBL; U36849; G1209648; JOINED.  
DR EMBL; U36850; G1209648; JOINED.  
DR EMBL; U36851; G1209648; JOINED.

DR EMBL; U36852; G1209648; JOINED.  
DR EMBL; U36853; G1209648; JOINED.  
DR EMBL; U36854; G1209648; JOINED.  
DR EMBL; U36855; G1209648; JOINED.  
DR EMBL; U36856; G1209648; JOINED.  
DR PIR; A36182; TDFELK.  
DR FLYBASE; FBgn0000484; Lar.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 2.  
DR PROSITE; PS00506; TYR\_PHOSPHATASE\_2; 2.  
DR PROSITE; PS00505; TYR\_PHOSPHATASE\_Ptp; 2.  
DR PFAM; PF00041; fn3; 9.  
DR PFAM; PF00047; 1g; 3.  
DR PFAM; PF00102; Y\_phosphatase; 2.  
DR HSSP; P28827; 1RPM.  
KW HYDROLASE; RECEPTOR; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW CELL ADHESION; IMMUNOGLOBULIN FOLD; DUPLICATION.  
FT SIGNAL 1 32  
FT CHAIN 33 2029 PROTEIN TYROSINE PHOSPHATASE DLAR.  
FT DOMAIN 33 1377 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1378 1402 POTENTIAL.  
FT DOMAIN 1403 2029 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 50 118 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 154 216 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 249 308 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 320 417 FIBRONECTIN TYPE-III.  
FT DOMAIN 418 512 FIBRONECTIN TYPE-III.  
FT DOMAIN 513 607 FIBRONECTIN TYPE-III.  
FT DOMAIN 608 706 FIBRONECTIN TYPE-III.  
FT DOMAIN 707 809 FIBRONECTIN TYPE-III.  
FT DOMAIN 810 906 FIBRONECTIN TYPE-III.  
FT DOMAIN 907 1007 FIBRONECTIN TYPE-III.  
FT DOMAIN 1008 1102 FIBRONECTIN TYPE-III.  
FT DOMAIN 1103 1207 FIBRONECTIN TYPE-III.  
FT DOMAIN 1492 1738 FIBRONECTIN TYPE-III.  
FT DOMAIN 1781 2029 FIBRONECTIN TYPE-III.  
FT ACT\_SITE 1670 1670 PROTEIN-TYROSINE PHOSPHATASE.  
FT ACT\_SITE 1961 1961 BY SIMILARITY.  
FT DISULFID 57 111 POTENTIAL.  
FT DISULFID 161 209 POTENTIAL.  
FT DISULFID 256 301 POTENTIAL.  
FT CARBOHYD 176 176 POTENTIAL.  
FT CARBOHYD 253 253 POTENTIAL.  
FT CARBOHYD 298 298 POTENTIAL.  
FT CARBOHYD 553 553 POTENTIAL.  
FT CARBOHYD 616 616 POTENTIAL.  
FT CARBOHYD 666 666 POTENTIAL.  
FT CARBOHYD 721 721 POTENTIAL.  
FT CARBOHYD 774 774 POTENTIAL.  
FT CARBOHYD 915 915 POTENTIAL.  
FT CARBOHYD 962 962 POTENTIAL.  
FT CARBOHYD 1183 1183 POTENTIAL.  
FT CARBOHYD 1304 1304 POTENTIAL.  
SQ SEQUENCE 2029 AA; 229027 MW; 6BDF8F93 CRC32;

Query Match 60.9%; Score 39; DB 1; Length 2029;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLKWKYK 9  
| | | | |  
Db 339 RLENSYK 346

RESULT 9  
YQR5\_YEAST  
ID YQR5\_YEAST STANDARD; PRT; 106 AA.  
AC P32339;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN IN ORIS 5'REGION (FRAGMENT).  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).





DE 60S RIBOSOMAL PROTEIN L10.  
GN RPL10.  
OS EUGLENA GRACILIS.  
RN EUKARYOTA; EUGLENOZOEA; EUGLENIDA; EUGLENALES; EUGLENA.  
RC [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Z;  
RA DERUERE J., SCHANTZ M.L., SCHANTZ R.;  
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: BELONGS TO THE L10E FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X94087; E213753; -  
DR PROSITE; PS01257; RIBOSOMAL\_L10E; 1.  
DR PFAM; PF00826; Ribosomal\_L10e; 1.  
KW RIBOSOMAL PROTEIN.  
SQ SEQUENCE 215 AA; 24642 MW; CBF0FC23 CRC32;  
  
Query Match 59.4%; Score 38; DB 1; Length 215;  
Best Local Similarity 66.7%; Pred. No. 8.2;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 3 LKWKYKGF 11  
: : : : :  
Db 182 VKWKEGKF 190  
  
RESULT 12  
UDAL\_RAT  
ID UDAL\_RAT STANDARD; PRT; 527 AA.  
AC P36510;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE UDP-GLUCURONOSYLTRANSFERASE 2A1 PRECURSOR, MICROSOMAL (EC 2.4.1.17)  
DE (UDEPT) (UGT-OLF).  
GN UGT2A1 OR UGT2A-1.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91156050.  
RA LAZARD D., ZUPKO K., PORIA Y., NEF P., LAZAROVITS J., HORN S.,  
RA KHEN M., LANCET D.;  
RT "Odorant signal termination by olfactory UDP glucuronosyl  
transferase";  
RL NATURE 349:790-793(1991).  
RN [4]  
CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
CC ENDOGENOUS COMPOUNDS. THIS ISOFORM IS ACTIVE ON ODORANTS AND  
CC SEEMS TO BE INVOLVED IN OLFACTION; IT COULD HELP CLEAR LIPOPHILIC  
CC ODORANT MOLECULES FROM THE SENSORY EPITHELIUM.  
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR  
CC BETA-D-GLUCORONOSIDE.  
CC -1- TISSUE SPECIFICITY: OLFACTORY EPITHELIUM.  
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL: X57565; -; NOT\_ANNOTATED\_CDS.  
DR PROSITE; PS00375; UDPGT; 1.  
DR PFAM; PF00201; UDPGT; 1.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL;  
KW MULTIGENE FAMILY; MICROSOME; OLFACTION.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 527 UDP-GLUCURONOSYLTRANSFERASE 2A1.  
FT TRANSHEM 491 507 POTENTIAL.  
FT CARBOHYD 49 49 POTENTIAL.  
FT CARBOHYD 313 313 POTENTIAL.  
SQ SEQUENCE 527 AA; 59915 MW; F81D3031 CRC32;  
  
Query Match 59.4%; Score 38; DB 1; Length 527;  
Best Local Similarity 50.0%; Pred. No. 21;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 KRLWKYKKGK 10  
: : : : :  
Db 331 QKVLWRYKKGK 340  
  
RESULT 13  
HXDB\_CHICK  
ID HXDB\_CHICK STANDARD; PRT; 280 AA.  
AC P24342; Q90780;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HOMEBOX PROTEIN HOX-D11 (CHOX-4.6) (CHOX-4.6) (CHOX-4E).  
GN HOXD11 OR CHOX-4.6.  
OS GALLUS GALLUS (CHICKEN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92173259.  
RA REGINA B., COELHO C.N., KOSHER R.A., UPHOLT W.B.;  
RT "The pattern of expression of the chicken homolog of HOX11 in the  
RT developing limb suggests a possible role in the ectodermal inhibition  
RT of chondrogenesis";  
RL DEV. DYN. 193:92-101(1992).  
RN [2]  
RP SEQUENCE OF 181-280 FROM N.A.  
RA NOHNO T.;  
RL SUBMITTED (FEB-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 208-268 FROM N.A.  
RX MEDLINE; 91168267.  
RA NOHNO T., NOJI S., KOYAMA E., OHYAMA K., MYOKAI F., KUROIWA A.,  
RA SAITO T., TANIGUCHI S.;  
RT "Involvement of the Chox-4 chicken homeobox genes in determination of  
RT anteroposterior axial polarity during limb development";  
RL CELL 64:1197-1205(1991).  
RN [4]  
RP SEQUENCE OF 208-267 FROM N.A.  
RX MEDLINE; 91204053.  
RA IZPISUA-BELMONT J.-C., TICKLE C., DOLLE P., WOLPERT L., DUBOULE D.;  
RT "Expression of the homeobox Hox-4 genes and the specification of  
RT position in chick wing development";  
RL NATURE 350:585-589(1991).  
RN [4]  
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF  
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH  
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- OVERLAPPING DOMAINS DURING WING DEVELOPMENT.  
CC -1- SIMILARITY: BELONGS TO THE ABD-B FAMILY OF HOMEBOX PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M81249; G211937; -;  
 DR EMBL: D10288; G222805; -;  
 DR PIR: B37914; B37914.  
 DR PIR: S14799; S14799.  
 DR PROSITE: PS00027; HOMEBOX.1; 1.  
 DR PFAM: PF00046; HOMEBOX.2; 1.  
 DR HSPF: F02835; 1F72.  
 DR TRANSFAC: T01759; -;  
 KW HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN;  
 FT TRANSCRIPTION REGULATION.  
 FT DNA\_BIND 208 267 HOMEBOX.  
 FT CONFLICT 181 200 VTTASGAASPPGEVAEAKSN -> SSSQSAVVFCEINOHFL  
 FT SLI (IN REF. 2).  
 SQ SEQUENCE 280 AA; 31923 MW; 016DC9ED CRC32;

Query Match 57.8%; Score 37; DB 1; Length 280;  
 Best Local Similarity 45.5%; Pred. No. 16;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KRLWKYKGF 11  
 :|::|::|:  
 Db 68 ERGKWYRGSY 78

RESULT 14  
 NU2M\_DASNO  
 ID NU2M\_DASNO STANDARD; PRT; 347 AA.  
 AC Q21326;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).  
 GN MTND2 OR ND2 OR NADH2.  
 OS DASYPUS NOVEMCINCTUS (NINE-BANDED ARMADILLO).  
 OG MITOCHONDRION  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC EDENTATA; DASYPODIDAE; DASYPUS.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 97357423.  
 RA ARNASON U., GULLBERG A., JANKE A.;  
 RT "Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates.";  
 RL MOL. BIOL. EVOL. 14:762-768(1997).  
 CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL INNER MEMBRANE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y11832; E311017; -;  
 DR PFAM: PF00361; oxidored.q1.1.  
 KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.  
 SQ SEQUENCE 347 AA; 39329 MW; 363847A5 CRC32;

Query Match 57.8%; Score 37; DB 1; Length 347;  
 Best Local Similarity 71.4%; Pred. No. 20;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 RLKWKYK 8  
 :|::|::|:  
 Db 312 KMKWKYK 318

RESULT 15  
 PSAB\_YERPE  
 ID PSAB\_YERPE STANDARD; PRT; 273 AA.  
 AC P31523; Q56979;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE CHAPERONE PROTEIN PSAB PRECURSOR.  
 OS PSAB.  
 GN YERSINIA PESTIS, AND YERSINIA PSEUDOTUBERCULOSIS.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 CC YERSINIA.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=Y.PESTIS;  
 RX MEDLINE: 93302507.  
 RA LINDLER L.E., TALL B.D.;  
 RT "Yersinia pestis ph 6 antigen forms fimbriae and is induced by intracellular association with macrophages.";  
 RL MOL. MICROBIOL. 8:311-324(1993).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=Y.PESTIS; STRAIN-EV76;  
 RC CHEREPANOV P.A.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=Y.PSEUDOTUBERCULOSIS; STRAIN-YPIII;  
 RX MEDLINE: 96294755.  
 RA YANG Y., MERRIAM J.J., MUELLER J., ISBERG R.R.;  
 RT "The psa locus is responsible for thermoinducible binding of Yersinia pseudotuberculosis to cultured cells.";  
 RL INFECT. IMMUN. 64:2483-2489(1996).  
 CC -!- FUNCTION: REQUIRED FOR THE BIOGENESIS OF THE PH6 ANTIGEN.  
 CC -!- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).  
 CC -!- SIMILARITY: STRONG, TO OTHER PERIPLASMIC PILUS CHAPERONES.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.  
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-23 IS THE INITIATOR.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M86713; G155489; ALT\_INIT.  
 DR EMBL: X97759; E242706; -;  
 DR EMBL: L76301; G1197044; -;  
 DR PIR: S27748; S27748.  
 DR PROSITE: PS00635; PILL\_CHAPERONE; 1.  
 DR PFAM: PF00345; pill\_assembly; 1.  
 KW CHAPERONE; PERIPLASMIC; SIGNAL; IMMUNOGLOBULIN FOLD.  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 273 CHAPERONE PROTEIN PSAB.  
 FT DISULFID 128 163 POTENTIAL.  
 FT CONFLICT 252 252 E -> D (IN REF. 2).  
 SQ SEQUENCE 273 AA; 30648 MW; 9B4ESD43 CRC32;

Query Match 57.8%; Score 37; DB 1; Length 273;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Oy 3 LKWKYKGF 11  
 :|::|::|:

Db 184 LTWRYKGN 192

Search completed: September 8, 1999, 00:43:02  
Job time: 2617 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:54:54 ; Search time 31.14 Seconds  
(without alignments)  
14.153 Million cell updates/sec

Title: US-09-124-280A-45  
Perfect score: 64  
Sequence: 1 KRLWKYKGF 11

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database : PIR\_60:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	102	2 A23931	anti-lipoplysacch
2	61	95.3	101	2 A27819	anti-lipoplysacch
3	45	70.3	668	2 A46013	coagulation factor
4	42	65.6	661	1 KFH013	coagulation factor
5	41	64.1	563	2 S11175	choline transport
6	40	62.5	2222	2 A36028	DNA-directed DNA p
7	40	62.5	277	2 A43783	homeotic protein N
8	39	60.9	2029	1 TDFL0K	protein-tyrosine-p
9	39	60.9	682	2 B70121	hypothetical prote
10	39	60.9	273	2 S48556	probable membrane
11	38	59.4	59	2 S39068	homeotic protein c
12	38	59.4	527	2 S15089	glucuronosyltransf
13	38	59.4	168	2 E69119	conserved hypothet
14	38	59.4	293	2 D48997	tumor surface anti
15	38	59.4	305	2 C48997	tumor surface anti
16	38	59.4	312	2 B48997	tumor surface anti
17	38	59.4	323	2 A48997	tumor surface anti
18	37	57.8	273	2 S32927	psaB protein precu
19	36	56.2	1246	1 TDRFLT	leukocyte common a
20	36	56.2	766	1 WZBEC5	infected cell prot
21	36	56.2	445	1 F71033	hypothetical prote
22	36	56.2	1087	2 I51552	platelet-derived g
23	36	56.2	287	2 A45854	leucocyte common a
24	36	56.2	295	2 H71085	probable MRP prote
25	36	56.2	398	2 E64913	hypothetical prote
26	36	56.2	253	2 D71682	cytochrome C1, hem
27	36	56.2	480	2 F69774	transposon-related
28	36	56.2	957	2 H69141	hypothetical prote
29	36	56.2	337	2 S53846	ribosomal protein
30	36	56.2	438	2 G71175	hypothetical prote
31	36	56.2	758	2 D71072	hypothetical prote
32	35	54.7	323	2 S16318	homeotic protein H
33	35	54.7	329	2 C69667	ribonucleoside-dip
34	35	54.7	337	2 S74850	hypothetical prote
35	35	54.7	168	2 S77282	hypothetical prote
36	35	54.7	226	2 S32215	hypothetical prote
37	35	54.7	195	2 B69927	ribonucleoside-dip
38	35	54.7	161	2 G64355	hypothetical prote
39	35	54.7	954	2 S57108	hypothetical prote

40 35 54.7 389 2 S64420 mitochondrial inne  
41 35 54.7 1246 2 S60954 probable membrane  
42 35 54.7 1190 2 S47536 SWH1 protein - yea  
43 35 54.7 737 2 A32156 UDPglucose--glyco  
44 35 54.7 735 2 A33369 UDPglucose--glyco  
45 35 54.7 558 2 F71175 hypothetical prote

ALIGNMENTS

RESULT 1  
A23931  
anti-lipoplysaccharide factor - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 18-Jun-1993  
C:Accession: A23931  
R:Kuregawa, J.; Miyata, T.; Ohtsubo, S.; Nakamura, T.; Morita, T.; Hayashida, H.; M1  
J. Biol. Chem. 261, 7357-7365, 1986  
A:Title: Primary structure of Limulus anticoagulant anti-lipoplysaccharide factor.  
A:Reference number: A23931; MUID:86224003  
A:Accession: A23931  
A:Molecule type: protein  
A:Residues: 1-102 <AKE>

Query Match 100.0%; Score 64; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.0003;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11  
|||||  
DB 41 KRLWKYKGF 51

RESULT 2  
A27819  
anti-lipoplysaccharide factor - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 18-Jun-1993  
C:Accession: A27819  
R:Muta, T.; Miyata, T.; Tokunaga, F.; Nakamura, T.; Iwanaga, S.  
J. Biochem. 101, 1321-1330, 1987  
A:Title: Primary structure of anti-lipoplysaccharide factor from American horseshoe  
A:Reference number: A27819; MUID:88032893  
A:Accession: A27819  
A:Molecule type: protein  
A:Residues: 1-101 <MUT>  
A:Note: 13-Lys was also found  
F:31-52/Disulfide bonds: #status experimental

Query Match 95.3%; Score 61; DB 2; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11  
:|||||  
DB 40 RRLWKYKGF 50

RESULT 3  
A46013  
coagulation factor XIII subunit b - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Dec-1998  
C:Accession: A46013  
R:Nonaka, M.; Matsuda, Y.; Shiroishi, T.; Moriaki, K.; Nonaka, M.; Natsume-Sakai, S  
Genomics 15, 535-542, 1993  
A:Title: Molecular cloning of the b subunit of mouse coagulation factor XIII and assl  
A:Reference number: A46013; MUID:93224141  
A:Accession: A46013  
A>Status: preliminary

A:Molecule type: nucleic acid  
 A:Residues: 1-668 <NON>  
 A:Cross-references: GB:D10071; NID:g303651; PID:d1001429; PID:g303652  
 A:Experimental source: liver  
 A:Note: sequence extracted from NCBI backbone (NCBIN:129405, NCBIP:129407)  
 C:Superfamily: coagulation factor XIII chain b; complement factor H repeat homology  
 F:25-87/Domain: complement factor H repeat homology <FH01>  
 F:91-146/Domain: complement factor H repeat homology <FH02>  
 F:153-208/Domain: complement factor H repeat homology <FH03>  
 F:213-267/Domain: complement factor H repeat homology <FH04>  
 F:274-327/Domain: complement factor H repeat homology <FH05>  
 F:336-389/Domain: complement factor H repeat homology <FH06>  
 F:396-450/Domain: complement factor H repeat homology <FH07>  
 F:454-515/Domain: complement factor H repeat homology <FH08>  
 F:582-646/Domain: complement factor H repeat homology <FH10>

Query Match 70.3%; Score 45; DB 2; Length 668;  
 Best Local Similarity 77.8%; Pred. No. 3.4;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLKWKYK 10  
 Db 467 QLKWKYEGK 475  
 :|||||:|

RESULT 4  
 KFHU13  
 coagulation factor XIII chain b precursor - human  
 N:Alternate names: fibrin-stabilizing factor; plasma transglutaminase b chain; protein-g  
 C:Species: Homo sapiens (man)  
 C:Date: 19-Apr-1991 #sequence\_revision 26-May-1994 #text\_change 20-Mar-1998  
 C:Accession: A36397; A23830; S09980  
 R:Bottemus, R.E.; Ichinose, A.; Davie, E.W.  
 Biochemistry 29, 11195-11209, 1990  
 A:Title: Nucleotide sequence of the gene for the b subunit of human factor XIII.  
 A:Reference number: A36397; MUID:91105054  
 A:Accession: A36397  
 A:Molecule type: DNA  
 A:Residues: 1-661 <BOT>  
 A:Cross-references: GB:J05294  
 R:Ichinose, A.; McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
 Biochemistry 25, 4633-4638, 1986  
 A:Title: Title: Amino acid sequence of the b subunit of human factor XIII, a protein com  
 A:Reference number: A23830; MUID:87026535  
 A:Accession: A23830  
 A:Molecule type: mRNA  
 A:Residues: 2-661 <ICH>  
 A:Cross-references: GB:M14057; NID:gl82838; PID:gl82839  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 R:Grundmann, U.; Nerlich, C.; Rein, T.; Zettlmeissl, G.  
 Nucleic Acids Res. 18, 2817-2818, 1990  
 A:Title: Complete cDNA sequence encoding the B subunit of human factor XIII.  
 A:Reference number: S09980; MUID:90251467  
 A:Accession: S09980  
 A:Molecule type: mRNA  
 A:Residues: 1-661 <GRU>  
 A:Cross-references: EMBL:X51823  
 C:Comment: This protein is noncatalytic.  
 C:Genetics:  
 A:Gene: GDB:F13B  
 A:Cross-references: GDB:119893; OMIM:134580  
 A:Map position: lq31-lq32.1  
 A:Introns: 2/1; 69/1; 131/1; 190/1; 249/1; 309/1; 432/1; 499/1; 560/1; 631/2  
 C:Complex: heterotetramer in plasma with coagulation factor XIII chain a, the plasma tra  
 C:Superfamily: coagulation factor XIII chain b; complement factor H repeat homology  
 C:Keywords: blood coagulation; duplication; glycoprotein; plasma  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-661/Product: coagulation factor XIII chain b #status predicted <MAT>  
 F:25-87/Domain: complement factor H repeat homology <FH01>  
 F:91-146/Domain: complement factor H repeat homology <FH02>  
 F:153-208/Domain: complement factor H repeat homology <FH03>  
 F:213-267/Domain: complement factor H repeat homology <FH04>

F:274-327/Domain: complement factor H repeat homology <FH05>  
 F:336-389/Domain: complement factor H repeat homology <FH06>  
 F:396-450/Domain: complement factor H repeat homology <FH07>  
 F:454-515/Domain: complement factor H repeat homology <FH08>  
 F:524-578/Domain: complement factor H repeat homology <FH09>  
 F:582-646/Domain: complement factor H repeat homology <FH10>  
 F:617-619/Region: cell attachment (R-G-D) motif  
 F:25-76, 59-87, 91-135, 118-146, 153-197, 180-208, 213-255, 241-267, 274-316, 302-327, 336-378,  
 F:162, 545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.6%; Score 42; DB 1; Length 661;  
 Best Local Similarity 75.0%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LKWKYK 10  
 Db 468 MKWKYEGK 475  
 :|||||:

RESULT 5  
 S11175  
 choline transport protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein G3213; protein YGL077c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-Dec-1997  
 C:Accession: S11175; S64084  
 R:Nikawa, J.I.; Hosaka, K.; Tsukagoshi, Y.; Yamashita, S.  
 J. Biol. Chem. 265, 15996-16003, 1990  
 A:Title: Primary structure of the yeast choline transport gene and regulation of its  
 A:Reference number: S11175; MUID:90368823  
 A:Accession: S11175  
 A:Molecule type: DNA  
 A:Residues: 1-563 <NIK>  
 A:Cross-references: EMBL:J05603; NID:gl171329; PID:gl171330  
 R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
 submitted to the Protein Sequence Database, May 1996  
 A:Reference number: S64071  
 A:Accession: S64084  
 A:Molecule type: DNA  
 A:Residues: 1-563 <RIE>  
 A:Cross-references: EMBL:272599; NID:gl322592; PID:e243965; PID:gl322593; MIPS:YGL077  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:HNM1; CTR1  
 A:Cross-references: MIPS:YGL077c; SGD:S0003045  
 A:Map position: 7L  
 C:Superfamily: choline transport protein  
 C:Keywords: transmembrane #status predicted <TM1>  
 F:91-107/Domain: transmembrane #status predicted <TM2>  
 F:185-201/Domain: transmembrane #status predicted <TM3>  
 F:214-230/Domain: transmembrane #status predicted <TM4>  
 F:237-273/Domain: transmembrane #status predicted <TM5>  
 F:298-314/Domain: transmembrane #status predicted <TM6>  
 F:345-361/Domain: transmembrane #status predicted <TM7>  
 F:401-417/Domain: transmembrane #status predicted <TM8>  
 F:428-444/Domain: transmembrane #status predicted <TM9>  
 F:466-482/Domain: transmembrane #status predicted <TM9>

Query Match 64.1%; Score 41; DB 2; Length 563;  
 Best Local Similarity 87.5%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LKWKYK 10  
 Db 508 LYWKYK 515  
 :|||||:

RESULT 6  
 A36028  
 DNA-directed DNA polymerase (EC 2.7.7.7) II catalytic chain - yeast (Saccharomyces ce  
 N:Alternate names: DNA-directed DNA polymerase II chain A; protein N0825; protein YNL

C:Species: Saccharomyces cerevisiae  
C:Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 17-Mar-1999  
C:Accession: A36028; B36028; S60919; S63235; S65121  
R:Morrison, A.; Araki, H.; Clark, A.B.; Hamatake, R.K.; Sugino, A.  
Cell 62, 1143-1151, 1990  
A:Title: A third essential DNA polymerase in Saccharomyces cerevisiae.  
A:Reference number: A36028; MUID:90381771  
A:Accession: A36028  
A:Molecule type: DNA  
A:Residues: 1-2222 <MOR>  
A:Cross-references: GB:M60416; GB:M36724; NID:g171408; PID:g171409  
A:Accession: B36028  
A:Molecule type: protein  
A:Residues: 1214-1216, X, 1218-1221 <MO2>  
R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV from  
A:Reference number: S60909  
A:Accession: S60919  
A:Molecule type: DNA  
A:Residues: 1-2221 <SEN>  
A:Cross-references: EMBL:X92494; NID:g1045236; PID:g1045247  
R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63235  
A:Accession: S63235  
A:Molecule type: DNA  
A:Residues: 1-2222 <SEW>  
A:Cross-references: EMBL:Z71538; NID:g1302316; PID:e239648; PID:g1302317; MIPS:YNL262w  
A:Experimental source: strain S288C  
R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.  
Yeast 12, 505-514, 1996  
A:Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa  
A:Reference number: S65111; MUID:96310631  
A:Accession: S65121  
A:Molecule type: DNA  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Residues: 1-2221 <SEF>  
A:Cross-references: EMBL:X92494; NID:g1045236; PID:g1045247  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
C:Genetics:  
A:Gene: SGD:POL2; DUN2  
A:Cross-references: SGD:S0005206; MIPS:YNL262w  
A:Map position: 14L  
C:Superfamily: DNA-directed DNA polymerase II  
C:Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger

Query Match 62.5%; Score 40; DB 2; Length 2222;  
Best Local Similarity 45.5%; Pred. No. 82;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
:::|::|::|::|  
Db 679 RKLKWKWGEF 689

RESULT 7  
A43783  
homeotic protein NrHbox2 - eastern newt  
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
C:Date: 11-Dec-1992 #sequence\_revision 11-Dec-1992 #text\_change 17-Oct-1997  
C:Accession: A43783  
R:Brown, R.; Brookes, J.P.  
Development 111, 489-496, 1991  
A:Title: Identification and expression of a regeneration-specific homeobox gene in the  
A:Reference number: A43783; MUID:91372146  
A:Accession: A43783  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-277 <BRO>  
A:Cross-references: GB:X57674  
C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:206-262/Domain: homeobox homology <HGX>

Query Match 62.5%; Score 40; DB 2; Length 277;  
Best Local Similarity 45.5%; Pred. No. 9.8;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
:::|::|::|::|  
Db 68 RRSKWQYRGSY 78

RESULT 8

TDFLTK

N:Alternate tyrosine-phosphatase (EC 3.1.3.48) DLAR precursor - fruit fly (Drosophila mel  
P:Alternate names: leukocyte antigen-related protein  
C:Species: Drosophila melanogaster  
C:Date: 14-Dec-1990 #sequence\_revision 02-May-1994 #text\_change 12-Feb-1999  
C:Accession: A36182  
R:Streuli, M.; Krueger, N.X.; Tsai, A.Y.M.; Saito, H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 8698-8702, 1989  
A:Title: A family of receptor-linked protein tyrosine phosphatases in humans and Dros  
A:Reference number: A36182; MUID:90046860  
A:Accession: A36182

A:Molecule type: mRNA  
A:Residues: 1-2029 <STR>  
A:Cross-references: GB:M27700; NID:g157811; PID:g157812  
C:Genetics:  
A:Gene: FlyBase:Lar  
A:Cross-references: FlyBase:FBgn0000464  
C:Superfamily: leukocyte antigen-related protein; fibronectin type III repeat homolog  
ogy  
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembra  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-2029/Product: leukocyte antigen-related protein #status predicted <MAT>  
F:33-1377/Domain: extracellular #status predicted <EXT>  
F:50-113/Domain: immunoglobulin homology <IMM1>  
F:154-211/Domain: immunoglobulin homology <IMM2>  
F:249-303/Domain: immunoglobulin homology <IMM3>  
F:321-401/Domain: fibronectin type III repeat homology <FN3A>  
F:416-502/Domain: fibronectin type III repeat homology <FN3B>  
F:514-599/Domain: fibronectin type III repeat homology <FN3C>  
F:610-699/Domain: fibronectin type III repeat homology <FN3D>  
F:708-802/Domain: fibronectin type III repeat homology <FN3E>  
F:811-896/Domain: fibronectin type III repeat homology <FN3F>  
F:903-993/Domain: fibronectin type III repeat homology <FN3G>  
F:1006-1091/Domain: fibronectin type III repeat homology <FN3H>  
F:1101-1198/Domain: fibronectin type III repeat homology <FN3I>  
F:1378-1402/Domain: transmembrane #status predicted <TMM>  
F:1403-2029/Domain: intracellular #status predicted <INT>  
F:1417-2029/Domain: leukocyte common antigen cytosolic domain homology <LAC>  
F:1457-1718/Domain: protein-tyrosine-phosphatase homology <PTP1>  
F:1786-2009/Domain: protein-tyrosine-phosphatase homology <PTP2>  
F:57-111,161-209,256-301/Disulfide bonds: #status predicted  
F:176/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:1670/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1676/Binding site: substrate phosphate (Arg) #status predicted  
F:1961/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1967/Binding site: substrate phosphate (Arg) #status predicted

Query Match 60.9%; Score 39; DB 1; Length 2029;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLLKWKYKGF 9  
:::|::|::|  
Db 339 RLEWSYKGF 346

RESULT 9

B70121

hypothetical protein BB0170 - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 05-Jun-1998  
 C:Accession: B70121  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: B70121  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-682 <KLE>  
 A:Cross-references: GB:AE001128; GB:AE000783; NID:g2688057; PID:g2688070; TIGR:BB0170  
 A:Experimental source: strain B31

Query Match 60.9%; Score 39; DB 2; Length 682;  
 Best Local Similarity 54.5%; Pred. No. 36;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11

DB 369 KRLKYEYSGEF 379

# RESULT 10

S48556  
 Probable membrane protein YLR205c - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein L8167.18  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 13-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 17-Mar-1999  
 C:Accession: S48556; S25342  
 R:Pauley, A.  
 submitted to the EMBL Data Library, September 1994  
 A:Description: The sequence of S. cerevisiae cosmid 8167.  
 A:Reference number: S48545  
 A:Accession: S48556  
 A:Molecule type: DNA  
 A:Residues: 1-273 <PAU>  
 A:Cross-references: EMBL:U14913; NID:g544497; PID:g544514; MIPS:YLR205c  
 R:Simon, M.; della Seta, F.; Sor, F.; Faye, G.  
 Yeast 8, 559-567, 1992  
 A:Title: Analysis of the MSS51 region on chromosome XII of Saccharomyces cerevisiae.  
 A:Reference number: S25342; MUID:92397593  
 A:Accession: S25342  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 78-83,'H',85-181,'RN' <SIM>  
 C:Genetics:  
 A:Map position: 12R  
 C:Keywords: transmembrane protein  
 F:105-121/Domain: transmembrane #status predicted <TM1>  
 F:249-265/Domain: transmembrane #status predicted <TM2>

Query Match 60.9%; Score 39; DB 2; Length 273;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11

DB 167 RLKWEYKKNY 176

# RESULT 11

S39068  
 homeotic protein cnx5 - Hydra magnipapillata (fragment)  
 C:Species: Hydra magnipapillata  
 C:Date: 07-Oct-1994 #sequence\_revision 30-Jan-1998 #text\_change 30-Jan-1998  
 C:Accession: S39068

R:Naito, M.; Ishiguro, H.; Fujisawa, T.; Kurosawa, Y.  
 FEBS Lett. 333, 271-274, 1993  
 A:Title: Presence of eight distinct homeobox-containing genes in cnidarians.  
 A:Reference number: S39066; MUID:94039774  
 A:Accession: S39068  
 A:Molecule type: nucleic acid  
 A:Residues: 1-59 <NAI>  
 A:Cross-references: EMBL:Z22640  
 C:Genetics:  
 A:Gene: cnx5  
 C:Superfamily: unassigned homeobox proteins; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:1-43/Domain: homeobox homology (fragment) <HOX>

Query Match 59.4%; Score 38; DB 2; Length 59;  
 Best Local Similarity 50.0%; Pred. No. 4.4;  
 Matches 7; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 1 KRLKWK---YK GK 10

DB 37 RRLKWKRVKGIK GK 50

# RESULT 12

SI5089  
 glucuronosyltransferase (EC 2.4.1.17) - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 13-Sep-1998  
 C:Accession: SI5089  
 R:Lazard, D.; Zupko, K.; Poria, Y.; Nef, P.; Lazarovits, J.; Horn, S.; Khen, M.; Lanc  
 Nature 349, 790-793, 1991  
 A:Title: Odorant signal termination by olfactory UDP glucuronosyl transferase.  
 A:Reference number: SI5089; MUID:91156050  
 A:Accession: SI5089  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <LAZ>  
 C:Superfamily: glucuronosyltransferase  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 59.4%; Score 38; DB 2; Length 527;  
 Best Local Similarity 50.0%; Pred. No. 41;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYK GK 10

DB 331 QKVLWRYK GK 340

# RESULT 13

B69119  
 conserved hypothetical protein MTH1892 - Methanobacterium thermoautotrophicum (strain  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
 C:Accession: B69119  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan,  
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
 A:Reference number: A69000; MUID:98037514  
 A:Accession: B69119  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-168 <MTH>  
 A:Cross-references: GB:AE000940; GB:AE000666; NID:g2623011; PID:g2623024  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1892

Query Match 59.4%; Score 38; DB 2; Length 168;



Search completed: September 7, 1999, 23:54:55  
Job time: 1906 sec

Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 LKWYKGF 11  
| | | | |  
Db 96 LSWKFRGEF 104

## RESULT 14

D48997  
tumor surface antigen OA3-293 - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Feb-1997  
C:Accession: D48997  
R:Campbell, I.G.; Freemont, P.S.; Foulkes, W.; Trowsdale, J.  
Cancer Res. 52, 5416-5420, 1992  
A:Title: An ovarian tumor marker with homology to vaccinia virus contains an IgV-like re  
A:Reference number: A48997; MUID:93007897  
A:Accession: D48997  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-293 <CAM>  
A:Experimental source: ovarian carcinoma line PE/01  
A:Note: sequence extracted from NCBI backbone (NCBIN:114647, NCBIP:114652)  
C:Keywords: surface antigen

Query Match 59.4%; Score 38; DB 2; Length 293;

Best Local Similarity 62.5%; Pred. No. 23;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LKWYKGF 10  
| | | | |  
Db 56 VKWFKGR 63

## RESULT 15

C48997  
tumor surface antigen OA3-305 - human  
N:Alternate names: integrin-associated protein  
C:Species: Homo sapiens (man)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Sep-1997  
C:Accession: C48997; S36644  
R:Campbell, I.G.; Freemont, P.S.; Foulkes, W.; Trowsdale, J.  
Cancer Res. 52, 5416-5420, 1992  
A:Title: An ovarian tumor marker with homology to vaccinia virus contains an IgV-like re  
A:Reference number: A48997; MUID:93007897  
A:Accession: C48997  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-305 <CAM>  
A:Experimental source: ovarian carcinoma line PE/01  
A:Note: sequence extracted from NCBI backbone (NCBIN:114645, NCBIP:114651)  
R:Lindberg, P.  
submitted to the EMBL Data Library, August 1993  
A:Reference number: S36644  
A:Accession: S36644  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-305 <LIN>  
A:Cross-references: EMBL:Z25521; NID:g396704; PID:g396705  
C:Keywords: surface antigen

Query Match 59.4%; Score 38; DB 2; Length 305;

Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LKWYKGF 10  
| | | | |  
Db 56 VKWFKGR 63

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:23:08 ; Search time 25.2 Seconds  
(without alignments)  
4.308 Million cell updates/sec

Title: US-09-124-280A-45  
Perfect score: 64  
Sequence: 1 KRLKWKYKGF 11

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	102	1	US-08-169-701-1
2	64	100.0	11	1	US-08-037-830E-28
3	64	100.0	11	1	US-08-456-112B-28
4	64	100.0	11	2	US-08-456-112B-45
5	64	100.0	102	3	PCT-US96-10227-1
6	61	95.3	101	1	US-08-264-244-1
7	61	95.3	105	1	US-08-264-244-3
8	61	95.3	101	1	US-08-478-689-1
9	61	95.3	105	1	US-08-478-689-3
10	61	95.3	101	1	US-08-476-940-1
11	61	95.3	105	1	US-08-476-940-3
12	61	95.3	101	1	US-08-704-872-1
13	61	95.3	105	1	US-08-704-872-3
14	61	95.3	101	2	US-08-577-464-2
15	59	92.2	11	2	US-08-366-953A-25
16	39	60.9	2004	1	US-08-375-709-15
17	39	60.9	2004	2	US-08-752-929-15
18	35	54.7	332	1	US-08-469-649-2
19	34	53.1	391	1	US-07-921-178A-2
20	34	53.1	25	1	US-07-921-178A-16
21	34	53.1	391	1	US-08-103-445-5
22	34	53.1	391	1	US-08-461-690B-5
23	34	53.1	548	1	US-08-468-700-37
24	34	53.1	548	2	US-08-645-971-5
25	34	53.1	548	2	US-08-488-220-35
26	34	53.1	548	2	US-08-468-698-35
27	34	53.1	389	2	US-08-501-003A-11
28	34	53.1	372	2	US-08-501-003A-12
29	34	53.1	391	2	US-08-501-003A-13
30	34	53.1	383	2	US-08-501-003A-14
31	34	53.1	398	2	US-08-501-003A-15
32	34	53.1	391	2	US-08-501-003A-16
33	34	53.1	548	3	PCT-US94-01553A-35
34	34	53.1	548	3	PCT-US95-10426-35
35	32	50.0	138	1	US-07-827-506B-4
36	32	50.0	510	1	US-08-097-829-2
37	32	50.0	426	1	US-08-336-583-2
38	32	50.0	510	2	US-08-577-403-2
39	32	50.0	376	2	US-08-709-923-4

40 32 50.0 426 3 PCT-US95-13795-2 Sequence 2, Appl1  
41 31 48.4 483 1 US-08-194-338-7 Sequence 7, Appl1  
42 31 48.4 997 1 US-08-232-540-1 Sequence 1, Appl1  
43 31 48.4 519 1 US-08-319-387-4 Sequence 4, Appl1  
44 31 48.4 1021 1 US-08-233-008A-2 Sequence 2, Appl1  
45 31 48.4 998 1 US-08-233-008A-6 Sequence 6, Appl1

ALIGNMENTS

RESULT 1  
US-08-169-701-1  
: Sequence 1, Application US/08169701  
: Patent No. 5545721  
: GENERAL INFORMATION:  
: APPLICANT: CARROLL, SEAN B.  
: APPLICANT: FIRCA, JOSEPH R.  
: APPLICANT: FUGH, CHARLES  
: APPLICANT: PADHYE, NISHA V.  
: TITLE OF INVENTION: PREVENTION AND TREATMENT OF SEPSIS  
: NUMBER OF SEQUENCES: 1  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: HAVERTOCK, MEDLEN & CARROLL  
: STREET: 220 MONTGOMERY STREET, SUITE 2200  
: CITY: SAN FRANCISCO  
: STATE: CALIFORNIA  
: COUNTRY: UNITED STATES OF AMERICA  
: ZIP: 94104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA: US/08/169,701  
: APPLICATION NUMBER: US/08/169,701  
: FILING DATE: 17-DEC-1993  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/995,388  
: FILING DATE: 21-DEC-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: CARROLL, PETER G.  
: REGISTRATION NUMBER: 32,837  
: REFERENCE/DOCKET NUMBER: OPHD-00748  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 705-8410  
: TELEFAX: (415) 397-8338  
: INFORMATION FOR SEQ ID NO: 1:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 102 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: unknown  
: MOLECULE TYPE: peptide  
: FEATURE:  
: NAME/KEY: Peptide  
: LOCATION: 36  
: OTHER INFORMATION: /note= "The amino acid at this  
: OTHER INFORMATION: location can be either Val or Ile."  
: FEATURE:  
: NAME/KEY: Peptide  
: LOCATION: 102  
: OTHER INFORMATION: /note= "The amino acid at this  
: OTHER INFORMATION: location can be either Gln or Glu."  
us-08-169-701-1

Query Match 100.0%; Score 64; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KRLKWKYKGF 11

Db 41 KRLKWKYKGF 51  
 |||||

## RESULT 2

US-08-097-830E-28  
 ; Sequence 28, Application US/08097830E  
 ; Patent No. 5632211  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: Peptides For Neutralizing The  
 ; TOXICITY OF LIPID A  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan, P.C.  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/097,830E  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-097-830E-28

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 |||||

## RESULT 3

US-08-456-112B-28  
 ; Sequence 28, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular  
 ; US-08-456-112B-28

Query Match 100.0%; Score 64; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 |||||

Db 1 KRLKWKYKGF 11

## RESULT 4

US-08-456-112B-45  
 ; Sequence 45, Application US/08456112B  
 ; Patent No. 5834430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Porro, Massimo  
 ; TITLE OF INVENTION: POTENTIATION OF ANTIBIOTICS  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hedman, Gibson & Costigan  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: LEADING EDGE 486  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/456,112B  
 ; FILING DATE: May 31, 1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Costigan, James V.  
 ; REGISTRATION NUMBER: 25,669  
 ; REFERENCE/DOCKET NUMBER: 576-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 302-8989  
 ; TELEFAX: (212) 302-8998  
 ; INFORMATION FOR SEQ ID NO: 45:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 11 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: circular

US-08-456-112B-45

Query Match 100.0%; Score 64; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00017; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
| | | | | | | | | | |  
Db 1 KRLKWKYKGF 11

RESULT 5

PCT-US96-10227-1  
; Sequence 1, Application PC/TUS9610227  
; GENERAL INFORMATION:  
; APPLICANT: WILLIAMS, JAMES A.  
; APPLICANT: SHEKHANI, MOHAMMED S.  
; APPLICANT: FIRCA, JOSEPH R.  
; APPLICANT: SCHATZ, ROBERT W.  
; APPLICANT: PUGH, CHARLES  
; APPLICANT: PANASIK JR., NICHOLAS  
; APPLICANT: STAFFORD, DOUGLAS C.  
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF SEPSIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10227  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/482,191  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/169,701  
; FILING DATE: 17-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/164,067  
; FILING DATE: 08-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/995,388  
; FILING DATE: 21-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: OPHD-01280  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 36  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Val or Ile."  
; FEATURE:  
; NAME/KEY: Peptide

; LOCATION: 102  
; OTHER INFORMATION: /note= "The amino acid at this  
; OTHER INFORMATION: location can be either Gln or Glu."  
PCT-US96-10227-1

Query Match 100.0%; Score 64; DB 3; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
| | | | | | | | | | |  
Db 41 KRLKWKYKGF 51

RESULT 6

US-08-264-244-1  
; Sequence 1, Application US/08264244  
; Patent No. 5594113  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5594113man R.  
; APPLICANT: No. 5594113itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/264,244  
; FILING DATE: 22-JUNE-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-264-244-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
| | | | | | | | | | |  
Db 40 RRLKWKYKGF 50

RESULT 7

US-08-264-244-3  
; Sequence 3, Application US/08264244  
; Patent No. 5594113  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5594113man R.  
; APPLICANT: No. 5594113itsky, Thomas J.

;; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
;; TITLE OF INVENTION: Protein and Uses thereof  
;; NUMBER OF SEQUENCES: 3  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
;; STREET: 1100 New York Avenue, N.W.  
;; CITY: Washington  
;; STATE: District of Columbia  
;; COUNTRY: United States of America  
;; ZIP: 20005-3934  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/264,244  
;; FILING DATE: 22-JUNE-1994  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Esmond, Robert W.  
;; REGISTRATION NUMBER: 32,893  
;; REFERENCE/DOCKET NUMBER: 1413.0010005  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; TELEX: 248636 SSK  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 105 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
US-08-264-244-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 44 RRLKWKYKGF 54

RESULT 8  
US-08-478-689-1  
; Sequence 1, Application US/08478689  
; Patent No. 5614369  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5614369man R.  
; APPLICANT: No. 5614369itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,689  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-478-689-1

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Esmond, Robert W.  
;; REGISTRATION NUMBER: 32,893  
;; REFERENCE/DOCKET NUMBER: 1413.0010005  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; TELEX: 248636 SSK  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 101 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
US-08-478-689-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 40 RRLKWKYKGF 50

RESULT 9  
US-08-478-689-3  
; Sequence 3, Application US/08478689  
; Patent No. 5614369  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5614369man R.  
; APPLICANT: No. 5614369itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,689  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-478-689-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGF 11  
:|||||  
Db 44 RRLWKYKGF 54

RESULT 10  
US-08-476-940-1  
; Sequence 1, Application US/08476940  
; Patent No. 5627266  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5627266man R.  
; APPLICANT: No. 5627266itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,940  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 101 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-476-940-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGF 11  
:|||||  
Db 40 RRLWKYKGF 50

RESULT 11  
US-08-476-940-3  
; Sequence 3, Application US/08476940  
; Patent No. 5627266  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5627266man R.  
; APPLICANT: No. 5627266itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.

; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,940  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1413.0010005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; TELEX: 248636 SSK  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-476-940-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGF 11  
:|||||  
Db 44 RRLWKYKGF 54

RESULT 12  
US-08-704-872-1  
; Sequence 1, Application US/08704872  
; Patent No. 5747455  
; GENERAL INFORMATION:  
; APPLICANT: Wainwright, No. 5747455man R.  
; APPLICANT: No. 5747455itsky, Thomas J.  
; TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
; TITLE OF INVENTION: Protein and Uses Thereof  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: District of Columbia  
; COUNTRY: United States of America  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,872  
; FILING DATE: 30-AUG-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/264,244  
; FILING DATE: 22-JUNE-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 1413.0010005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-704-872-1

Query Match 95.3%; Score 61; DB 1; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
DB 40 RRLKWKYKGF 50

RESULT 13  
US-08-704-872-3  
Sequence 3, Application US/08704872  
Patent No. 5747455  
GENERAL INFORMATION:  
APPLICANT: Wainwright, No. 5747455man R.  
APPLICANT: No. 5747455itsky, Thomas J.  
TITLE OF INVENTION: Endotoxin Binding and Neutralizing  
TITLE OF INVENTION: Protein and Uses Thereof  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: United States of America  
ZIP: 20005-3934

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704.872  
FILING DATE: 30-AUG-1996  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/264,244  
FILING DATE: 22-JUNE-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1413.0010005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
TELEX: 248636 SSK  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 105 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-704-872-3

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0035;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||

Db 44 RRLKWKYKGF 54  
RESULT 14  
US-08-577-464-2  
Sequence 2, Application US/08577464  
Patent No. 5780429  
GENERAL INFORMATION:  
APPLICANT: WAINWRIGHT, No. 5780429man R.  
TITLE OF INVENTION: ANTI-LPS FACTOR FROM HORSESHOE CRABS AND  
TITLE OF INVENTION: METHODS OF USE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorrill  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/577,464  
FILING DATE: 22-DEC-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: BAKER, HOLLIE L.  
REGISTRATION NUMBER: 31,321  
REFERENCE/DOCKET NUMBER: 58378.650  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-942-8400  
TELEFAX: 202-942-8484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 101 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-577-464-2

Query Match 95.3%; Score 61; DB 2; Length 101;  
Best Local Similarity 90.9%; Pred. No. 0.0034;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
:|||||  
Db 40 RRLKWKYKGF 50

RESULT 15  
US-08-366-953A-25  
Sequence 25, Application US/08366953A  
Patent No. 5766593  
GENERAL INFORMATION:  
APPLICANT: Lichenstein, Henri S.  
APPLICANT: Wright, Samuel D.  
TITLE OF INVENTION: ANTI-INFLAMMATORY CD14 PEPTIDES  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMGEN INC.  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



;  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/366,953A  
; FILING DATE: 30-DEC-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cook Ph.D., Robert R.  
; REGISTRATION NUMBER: 31,602  
; REFERENCE/DOCKET NUMBER: A-324  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-366-953A-25

Query Match 92.2%; Score 59; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00099;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11  
Db 1 RLKWKYKGF 10

Search completed: September 7, 1999, 23:23:09  
Job time: 913 sec

**THIS PAGE BLANK (USPTO)**

Accession	Gene	Accession	Gene
44	34 53.1 431 1 W38420	Human multiple mye	
44	34 53.1 431 1 W38420	Human Prt1-like su	
45	34 53.1 873 1 W49031		

## ALIGNMENTS

Accession	Gene	Accession	Gene	Accession	Gene
44	Human multiple myeloma	53.1	451	1	W38426
45	Human Prt1-like su	53.1	873	1	W49031

## RESULT 1

ID R71799 standard; peptide; 11 AA.

DT 01-OCT-1995 (first entry)  
DE Peptide neutralising toxicity of Lipid A.  
KW Endotoxin; Lipid A; septic shock; lipopolysaccharide.

DE. Pentide neut

KW Endotoxin; Lipid A; sep

Us synthetic.

PD 02-FEB-1995.  
DE 21-THY-1994: F0241

PR 26-JUL-1993; US-09

PI PORRO M;

DR WPI; 95-075190/10.  
PT New peptide(s) for

repeating units of

PS Claim 30; Page 22;

CC Include units of f

CC of Arg and n is /- amino acid val. Ie

$$CC(ABC)^p, \text{ where } A \text{ is}$$

CC The peptides bind

activity. Hence the  
and also in vitro

CC nutrient solutions  
CC The present season

peptides.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Query Match

Best Local Similarity  
Matches 11: 100%

CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 |||||  
 Db 1 KRLKWKYKGF 11

RESULT 3  
 W21616 ID W21616 standard; peptide; 11 AA.  
 AC W21616;  
 DT 26-AUG-1997 (first entry)  
 DE Antibiotic potentiating peptide #28.  
 KW Potentiate; antibiotic; microbial infection; lipopolysaccharide;  
 KW permeability; outer bacterial membrane.  
 OS Synthetic.  
 PN WO9638163-A1.  
 PD 05-DEC-1996.  
 PF 29-MAY-1996; E02313.  
 PR 31-MAY-1995; US-456112.  
 PA (BIOS-) BIOSYNTH SRL.  
 PI Porro M, Varra M;  
 DR WPI: 97-034095/03.  
 PT Potentiating activity of antibiotic with peptide contg. cationic  
 PT amino acid sequence - reduces dose of antibiotic required  
 FS Claim 32; Page 27; 37pp; English.  
 CC The sequences given in W21589-633 represent peptides which act to  
 CC potentiate the activity of an antibiotic when they are co-administered  
 CC with the antibiotic. Compositions containing these peptides are used  
 CC to treat or prevent microbial infections. These peptides bind to  
 CC lipopolysaccharide on the bacteria so may increase permeability of  
 CC the outer bacterial membrane to the antibiotic, allowing a reduction  
 CC in the dose of antibiotic required by 10-90% of the normal dose for  
 CC in vivo or in vitro application. Any toxic side effects are  
 CC correspondingly reduced.  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 64; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 |||||  
 Db 1 KRLKWKYKGF 11

RESULT 4  
 W10350 ID W10350 standard; peptide; 102 AA.  
 AC W10350;  
 DT 22-SEP-1997 (first entry)  
 DE Limulus antilipopolysaccharide factor.  
 KW LALF; antimicrobial; immunoglobulin; surface binding ligand;  
 KW hapten; sepsis; influenza; viremia; fungemia; neurology; cancer;  
 KW endocrinology; antidiabetic; antibody; antibiotic.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT misc\_difference 36  
 FT /label- val, ile  
 PN WO9640251-A1.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; U10227.

PR 07-JUN-1995; US-482191.  
 PA (OPHI-) OPHIDIAN PHARM INC.  
 PI Firca JR, Panasiak N, Pugh C, Schatz RW, Shekhani MS;  
 PI Stafford D, Williams JA;  
 DR WPI: 97-077224/07.  
 PT Antimicrobial compns. for diagnosis and therapy - comprising  
 PT microbial surface binding ligand-spacer-hapten, non-covalently  
 PT associated with immunoglobulin  
 PS Example 38; Fig 13; 227pp; English.  
 CC The present sequence is Limulus antilipopolysaccharide factor  
 CC (LALF), which is a single chain peptide known to bind and neutralise  
 CC endotoxin. This is conjugated to human immunoglobulin IgG to produce  
 CC an antibody, which is a conjugated antibody and antibiotic. Making  
 CC a conjugate, comprises covalently linking a surface-binding ligand  
 CC (SBL) to a hapten via a spacer to form a ligand-spacer-hapten (LSH)  
 CC compound, and reacting the compound with immunoglobulin (Ig) under  
 CC conditions that allow for non-covalent binding of the Ig to the  
 CC hapten of the compound. The conjugate can be used for the diagnosis,  
 CC prevention and treatment of microbial infections, e.g. sepsis,  
 CC influenza, viremia or fungemia. They can also be used in neurology,  
 CC cancer and endocrinology, where Ig targeting can provide desired  
 CC therapeutic effects. The methods for producing a conjugate can use  
 CC small molecule targeting ligands to efficiently direct otherwise  
 CC unreactive Ig to microbial targets. This targeting allows Ig to react  
 CC with microbial structures that may not normally be accessible to Ig,  
 CC or are incapable of stimulating antibody production.  
 SQ Sequence 102 AA;

Query Match 100.0%; Score 64; DB 1; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 0.0017;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 |||||  
 Db 41 KRLKWKYKGF 51

RESULT 5  
 R29467 ID R29467 standard; Protein; 105 AA.  
 AC R29467;  
 DT 15-APR-1993 (first entry)  
 DE Modified endotoxin binding/neutralising protein.  
 KW Endotoxin; binding; neutralising; protein; horseshoe; crab; lysate;  
 KW cellular debris; limulus; amebocyte; hypotonic shock; yeast; urea;  
 KW denaturant; guanidine hydrochloride; tetrapeptide.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT peptide 1..4 /note= "Amino terminal modifying peptide"  
 FT PN W09220715-A.  
 PD 26-NOV-1992.  
 PF 15-MAY-1992; U03983.  
 PR 16-MAY-1991; US-701501.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Weinwright NR;  
 DR WPI: 92-415712/50.  
 DR N-PSDB; Q31793.  
 PT Use of endotoxin binding-neutralising protein isolated from  
 PT horseshoe crab - for reversing the effects of endotoxin in vivo,  
 PT and treating Gram negative bacterial infections, septicaemia,  
 PT toxic shock etc.  
 PS Claim 18; Fig 16; 67pp; English.  
 CC The sequence given is an endotoxin binding/neutralising protein  
 CC which may be isolated from any of the four known species of horseshoe  
 CC crab. This protein may be obtained from the cellular debris produced  
 CC during lysate production from Limulus amebocytes. The amebocytes are  
 CC lysed by hypotonic shock. The endotoxin binding/neutralising protein  
 CC was extracted from the cellular debris by using denaturants such as  
 CC urea and guanidine hydrochloride. The endotoxin binding/neutralisation  
 CC domain was found to occur between amino acids 34-59 of the sequence  
 CC given. The gene encoding this protein was isolated from a yeast host

CC and has the additional coding sequence between bases 1-12 which encodes  
 CC the tetrapeptide "Glu-Ala-Glu-Ala" indicated in the features table.  
 CC This peptide is covalently linked to the amino terminal of the  
 CC endotoxin binding/neutralising protein.  
 SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
 Best Local Similarity 90.9%; Pred. No. 0.0052;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||  
 Db 44 RRLKWKYKGF 54

## RESULT 6

ID R29468 standard; Protein; 101 AA.  
 AC R29468;  
 DT 15-APR-1993 (first entry)  
 DE Endotoxin binding/neutralising protein.  
 KW Endotoxin; binding; neutralising; protein; horseshoe; crab; lysate;  
 KW cellular debris; Limulus; amebocyte; hypotonic shock; yeast; urea;  
 KW denaturant; guanidine hydrochloride; tetrapeptide.  
 OS Limulus polyphemus.  
 PN W09220715-A.  
 PD 26-NOV-1992.  
 PF 15-MAY-1992; U03983.  
 PR 16-MAY-1991; US-701501.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Wainwright NR;  
 DR WPI; 92-415712/50.  
 PT Use of endotoxin binding-neutralising protein isolated from  
 PT horseshoe crab - for reversing the effects of endotoxin in vivo,  
 PT and treating Gram negative bacterial infections, septicaemia,  
 PT toxic shock etc.  
 PS Claim 17; Fig 15; 67pp; English.  
 CC The sequence given is an endotoxin binding/neutralising protein  
 CC which may be isolated from any of the four known species of horseshoe  
 CC crab. This protein may be obtained from the cellular debris produced  
 CC during lysate production from Limulus amebocytes. The amebocytes are  
 CC lysed by hypotonic shock. The endotoxin binding/neutralising protein  
 CC was extracted from the cellular debris by using denaturants such as  
 CC urea and guanidine hydrochloride. The endotoxin binding/neutralisation  
 CC domain was found to occur between amino acids 30-55 of the sequence  
 CC given.  
 SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||  
 Db 40 RRLKWKYKGF 50

## RESULT 7

R90148 standard; peptide; 22 AA.  
 AC R90148;  
 DT 06-MAR-1996 (first entry)  
 DE LALF-(31-52) peptide.  
 KW lipopolysaccharide-binding peptide; LBP; LPS; antibacterial; sepsis;  
 KW Limulus anti-LPS factor; LALF; endotoxin; amphipathic loop structure.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..22  
 FT domain /note= "loop-stabilising bond"  
 FT 8..15  
 FT PI /note= "LPS-binding domain"

PN W09505393-A2.  
 PD 23-FEB-1995.  
 PF 18-AUG-1994; E02747.  
 PR 18-AUG-1993; US-108415.  
 PR 25-AUG-1993; US-111625.  
 PA (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.  
 PI Hoess A, Liddington RC;  
 DR WPI; 95-098719/13.  
 PT New lipo:polysaccharide binding and neutralising peptide(s) - used  
 PT for treating or preventing LPS associated conditions or for  
 PT detecting or removing bacterial LPS.  
 PS Claim 11; Page 27; 46pp; English.  
 CC New LPS-binding peptides are claimed which have a domain of 8 amino  
 CC acids (AA) comprising (from N to C terminal): (1) a polar or positively  
 CC charged AA (pref. C, H, K, N, Q, R, S, T, W or Y); (2) a hydrophobic AA (pref.  
 CC A, F, H, I, L, M, V or W); (3) a basic AA (pref. H, K or R); (4) a hydrophobic  
 CC or positively charged AA (pref. A, F, H, I, K, L, M, R, V or W); (5) a  
 CC hydrophobic, polar or positively charged AA (pref. A, C, F, H, I, K, L, M, N, Q,  
 CC R, S, T, V, W or Y); (6) a positively charged AA (pref. K or R); (7) a  
 CC hydrophobic, polar or positively charged AA (pref. A, C, F, H, I, K, L, M, N, Q,  
 CC R, S, T, V, W or Y); and (8) a hydrophobic or positively charged AA (pref.  
 CC A, F, H, I, K, L, M, R, V or W). The peptide can alternatively be a  
 CC corresponding inverse sequence, or it can be an LPS-binding variation  
 CC of either sequence. The peptides have been designed by studying an  
 CC amphipathic loop structure found in the LPS-binding proteins Limulus  
 CC anti-LPS factor (LALF), rabbit and human lipopolysaccharide-binding  
 CC protein (LBP) and human bactericidal/permeability-increasing protein  
 CC (BPI). The present sequence is a preferred example of the new peptides.  
 CC The peptides bind with high affinity to endotoxin and are useful for  
 CC prevention and treatment of Gram-negative and Gram-positive bacterial  
 CC sepsis and for the treatment of bacterial and fungal infections, as  
 CC well as for neutralising effects associated with heparin such as  
 CC anti-coagulation, angiogenesis, and growth factor-induced tumour and  
 CC endothelial cell proliferation.  
 SQ Sequence 22 AA;

Query Match 95.3%; Score 61; DB 1; Length 22;  
 Best Local Similarity 90.9%; Pred. No. 0.0012;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11  
 :|||||  
 Db 10 RRLKWKYKGF 20

## RESULT 8

W13262 standard; Protein; 105 AA.  
 ID W13262;  
 DT 29-MAY-1997 (first entry)  
 DE Modified Limulus polyphemus endotoxin binding protein.  
 DE Modified; recombinant; horseshoe crab; endotoxin; binding protein;  
 KW neutralisation; in vivo; assay; treatment; septicaemia; yeast;  
 KW toxic shock; endotoxin related arthritis; gonorrhea; expression;  
 KW periodontal disease; spinal meningitis; amniotic fluid infection;  
 KW gram negative; bacterium; bacteria; infection; host.  
 OS Limulus polyphemus.  
 FH Key Location/Qualifiers  
 FT peptide 1..4  
 FT /note= "tetrapeptide to enable expression as  
 FT glycoprotein in yeast host"  
 PN US5594113-A.  
 PD 14-JAN-1997.  
 PF 23-JUN-1988; 210575.  
 PR 23-JUN-1988; US-210575.  
 PR 16-FEB-1990; US-480957.  
 PR 16-MAY-1991; US-701501.  
 PR 15-MAY-1992; US-883457.  
 PR 22-JUN-1994; US-264244.  
 PA (ASCA-) ASSOC CAPE COD INC.  
 PI Novitsky TJ, Wainwright NR;  
 DR WPI; 97-107149/10.

DR N-PSDB; T62063.  
PT New horse-shoe crab endotoxin-binding protein - useful for  
PT neutralising the effects of the endotoxin in an animal  
PS Claim 1: Columns 19-20: 29pp: English.  
CC The present sequence is a Limulus polyphemus (horseshoe crab)  
CC endotoxin binding protein modified for expression in a yeast host,  
CC which can be used to bind and neutralise endotoxin in vivo and in  
CC endotoxin assays. The protein can be used to treat septicemia,  
CC toxic shock, endotoxin related arthritis, gonorrhea, periodontal  
CC disease, spinal meningitis, amniotic fluid infection and gram  
CC negative bacterial infections in humans or other mammals. The  
CC protein has a very high binding constant, and its low molecular  
CC weight makes it likely to be less antigenic than anti-endotoxin  
CC antibodies.  
CC Fifty ng/kg of E. coli endotoxin was injected into 9 rabbits, and  
CC 15 minutes later 3 were injected with 5 microg of the protein, 3  
CC with 50 microg of the protein and 3 with PBS only. The control  
CC animals that received endotoxin and PBS only showed normal peak  
CC fever response 1 hour after toxin administration, while animals  
CC that received the protein showed a greatly reduced fever response  
CC proportional to the amount of protein administered.  
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLKWKYKGF 11  
DB 44 RRLKWKYKGF 54  
:|||||

RESULT 9  
W18326  
ID W18326 standard; Protein: 105 AA.  
AC W18326;  
DT 22-JUL-1997 (first entry)  
DE Endotoxin binding/neutralising protein.  
KW Horseshoe crab; L. polyphemus; endotoxin binding/neutralising protein;  
KW assay; water purity; cleanliness; pharmaceutical manufacture;  
KW kidney dialysis unit; Limulus amoebocyte lysate; LAL assay;  
KW septicemia; toxic shock; endotoxin related arthritis; gonorrhea;  
KW periodontal disease; spinal meningitis; amniotic fluid infection;  
KW septic shock; bacterial endotoxin.  
OS Limulus polyphemus.  
FH Key Location/Qualifiers  
FT peptide 1..4  
FT protein /note= "N-terminal extension"  
FT protein 5..105  
FT protein /note= "Endotoxin binding/neutralising protein"  
US5614369-A.  
PN 23-JUN-1988; 210575.  
PD 25-MAR-1997.  
PF 23-JUN-1988; US-210575.  
PR 16-FEB-1990; US-480957.  
PR 16-MAY-1991; US-701501.  
PR 15-MAY-1992; US-883457.  
PR 22-JUN-1994; US-264244.  
PR 07-JUN-1995; US-478689.  
PA (ASCA-) ASSOC CAPE COD INC.  
PI Novitsky TJ, Wainwright NR;  
DR WPI: 97-201472/18.  
DR N-PSDB; T68929.  
PT Assaying endotoxin using binding protein from horseshoe crab - and  
PT measuring quenching of tryptophan fluorescence caused by endotoxin  
PT binding  
PS Claim 1: Column 19-20: 29pp: English.  
CC This sequence represents the horseshoe crab (L. polyphemus) endotoxin  
CC binding/neutralising protein. This protein was used in the method of  
CC the invention for assaying endotoxin concentration. The method can  
CC be used to determine water purity and process cleanliness during  
CC pharmaceutical manufacture, to check kidney dialysis units and

CC generally wherever the Limulus amoebocyte lysate (LAL) assay is  
CC currently used. Also the protein, or a truncated version lacking the  
CC 1st 4 amino acids, can be used therapeutically to bind/neutralise  
CC endotoxin in vivo (e.g. in cases of septicemia, toxic shock, endotoxin  
CC related arthritis, gonorrhea, periodontal disease, spinal meningitis  
CC and amniotic fluid infection), and to remove endotoxin from solutions  
CC (e.g. extracorporeal treatment of septic shock). This protein has a  
CC very high binding constant for a wide range of bacterial endotoxins,  
CC and being of low molecular weight is likely to be less immunogenic  
CC than anti-endotoxin antibodies. The protein also avoids the multiple  
CC interferences associated with the conventional LAL assay.  
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLKWKYKGF 11  
DB 44 RRLKWKYKGF 54  
:|||||

RESULT 10  
W19248  
ID W19248 standard; Protein: 105 AA.  
AC W19248;  
DT 18-AUG-1997 (first entry)  
DE Modified horseshoe crab endotoxin binding/neutralising protein.  
KW Modified; horseshoe crab; endotoxin; binding; neutralisation;  
KW amino-terminal; tetrapeptide; yeast; restriction site; solid;  
KW resin; immobilisation; support; chromatographic; membrane;  
KW phase; latex microsphere; removal; solution; pharmaceutical;  
KW extracorporeal; treatment; septic shock; reagent; assay;  
KW fluorescence quenching; biosensor; determination;  
KW agglutination.  
OS Limulus polyphemus.  
FH Key Location/Qualifiers  
FT region 1..4  
FT protein /note= "additional amino-terminal tetrapeptide"  
US5627266-A.  
PN 06-MAY-1997.  
PD 23-JUN-1988; 210575.  
PF 23-JUN-1988; US-210575.  
PR 16-FEB-1990; US-480957.  
PR 16-MAY-1991; US-701501.  
PR 15-MAY-1992; US-883457.  
PR 22-JUN-1994; US-264244.  
PR 07-JUN-1995; US-476940.  
PA (ASCA-) ASSOC CAPE COD INC.  
PI Novitsky TJ, Wainwright NR;  
DR WPI: 97-271365/24.  
DR N-PSDB; T69763.  
PT Immobilised endotoxin binding protein - for removal or determination  
PT of endotoxin  
PS Claim 1: Columns 19-20: 29pp: English.  
CC The present sequence is a modified Horseshoe crab endotoxin  
CC binding/neutralising protein having an additional amino-terminal  
CC tetrapeptide, the codons for which are yeast preferred codons  
CC possessing unique restriction sites. When immobilised on a solid  
CC support (preferably a chromatographic resin or membrane, or latex  
CC microspheres) the protein can be used to remove endotoxin from  
CC solutions, e.g. pharmaceutical solutions, for the extracorporeal  
CC treatment of septic shock, as a solid phase reagent in a fluorescence  
CC quenching assay for endotoxin, in biosensors for endotoxin  
CC determination or as an agglutination reagent for endotoxin  
CC determination.  
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
Best Local Similarity 90.9%; Pred. No. 0.0052;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 KRLWKYKGRF 11
   :|||||
Db 44 RRLWKYKGRF 54

RESULT 11
W18220
ID W18220 standard; Protein; 101 AA.
AC W18220;
DT 31-JAN-1998 (first entry)
DE Limulus anti-lipopolysaccharide factor.
KW Anti-lipopolysaccharide factor; LPS; Limulus anti-LPS factor;
KW LALF; Gram-negative bacterium; Infection; therapy; acne;
KW antimicrobial; preservative.
OS Limulus polyphemus.
OS Synthetic.
FH Key Location/Qualifiers
FT Domain 30..55
FT Note= "domain required for endotoxin binding and
   neutralisation"
FT W09723235-A1.
PD 03-JUL-1997.
PD 17-DEC-1996; U19882.
PD 22-DEC-1995; US-577464.
PA (MARI-) MARINE BIOLOGICAL LAB.
PI Wainwright NR;
DR WPI; 97-350781/32.
DR N-PSDB; T71720.
PT Anti-lipopolysaccharide factor proteins for use in cosmetic, skin or
PT hair preparations - useful to treat Gram-negative bacterial
PT infections, especially acne.
PS Example 1; Fig 1; 30pp; English.
CC This polypeptide comprises Limulus anti-lipopolysaccharide factor
CC (LALF) (W18220). It was expressed in Pichia host cells using a
CC synthetic gene (see T71720) that utilises Saccharomyces cerevisiae-
CC preferred codons. The yeast alpha-factor prepro leader sequence,
CC present on the expressed protein as an N-terminal fusion, targets
CC the protein for secretion into the medium and is cleaved off in the
CC process. A claimed pharmaceutical composition contains one or more
CC anti-lipopolysaccharide (anti-LPS) factor proteins and one or more
CC antibiotics. Also claimed is a preservative for use in cosmetic or
CC skin or hair preparations, comprising one or more anti-LPS factor
CC proteins. The composition is used to treat bacterial infections
CC (in conjunction with antimicrobics) and yeast infections. In
CC particular the composition is for treating skin acne, optionally in
CC conjunction with antibiotics.
SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;
Best Local Similarity 90.9%; Pred. No. 0.0051;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGRF 11
   :|||||
Db 40 RRLWKYKGRF 50

RESULT 12
W53876
ID W53876 standard; protein; 105 AA.
AC W53876;
DT 13-JUL-1998 (first entry)
DE L. polyphemus endotoxin.
KW Endotoxin contamination; horseshoe crab; gram negative septic shock;
KW bacterial infection.
OS Limulus polyphemus.
PN US5747455-A.
PD 05-MAY-1998.
PD 30-AUG-1996; 704872.
PR 15-MAY-1992; US-883457.
PR 23-JUN-1988; US-210575.
PR 16-FEB-1990; US-480957.
PR 16-FEB-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ Wainwright NR;
DR WPI; 98-285799/25.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Disclosure; Fig 15; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the

```

```

PR 16-FEB-1990; US-480957.
PR 16-MAY-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ Wainwright NR;
DR WPI; 98-285799/25.
DR N-PSDB; V23553.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Claim 1; Fig 16; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the
CC binding molecule; and (c) removing the complex for the sample. The method
CC is useful removing endotoxins from contaminated blood. Endotoxins are
CC lipid, carbohydrate and protein complexes from bacteria, normally found
CC in their outer wall. They are shed from living gram negative bacteria and
CC released when they die leading to gram negative septic shock from
CC bacterial infection, especially after surgery. The protein used in the
CC method binds to endotoxins and neutralises them. The method removes
CC endotoxins from infected patients. Previous antibiotic treatment of
CC bacterial infection only kills the bacteria, still leaving the endotoxin
CC in the system. Antibodies with the same purpose have been developed, but
CC the binding protein used has a higher binding constant and as it has a
CC low molecular weight, it is less antigenic than other proteins used in
CC similar procedures. The method is better at removing endotoxin than
CC previous methods e.g. ultrafiltration, ion exchange, affinity
CC chromatography etc. which cannot distinguish between proteins and
CC endotoxins.
SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;
Best Local Similarity 90.9%; Pred. No. 0.0052;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLWKYKGRF 11
   :|||||
Db 44 RRLWKYKGRF 54

RESULT 13
W53877
ID W53877 standard; protein; 101 AA.
AC W53877;
DT 13-JUL-1998 (first entry)
DE L. polyphemus endotoxin.
KW Endotoxin contamination; horseshoe crab; gram negative septic shock;
KW bacterial infection.
OS Limulus polyphemus.
PN US5747455-A.
PD 05-MAY-1998.
PD 30-AUG-1996; 704872.
PR 15-MAY-1992; US-883457.
PR 23-JUN-1988; US-210575.
PR 16-FEB-1990; US-480957.
PR 16-FEB-1991; US-701501.
PR 22-JUN-1994; US-264244.
PR 30-AUG-1996; US-704872.
PA (ASCA-) ASSOC CAPE COD INC.
PI Novitsky TJ Wainwright NR;
DR WPI; 98-285799/25.
PT Method for removing endotoxin contaminants - uses an endotoxin
PT binding protein from horseshoe crab
PS Disclosure; Fig 15; 29pp; English.
CC This sequence represents the Limulus polyphemus (horseshoe crab)
CC endotoxin. It can be used in the method of the invention for reducing
CC endotoxin contamination of a material containing it, comprising:
CC (a) contacting the material with the endotoxin binding molecule from
CC L. polyphemus; (b) forming a complex between the endotoxin and the

```

CC binding molecule; and (c) removing the complex for the sample. The method  
 CC is useful removing endotoxins from contaminated blood. Endotoxins are  
 CC lipid, carbohydrate and protein complexes from bacteria, normally found  
 CC in their outer wall. They are shed from living gram negative bacteria and  
 CC released when they die leading to gram negative septic shock from  
 CC bacterial infection, especially after surgery. The protein used in the  
 CC method binds to endotoxins and neutralises them. The method removes  
 CC endotoxins from infected patients. Previous antibiotic treatment of  
 CC bacterial infection only kills the bacteria, still leaving the endotoxin  
 CC in the system. Antibodies with the same purpose have been developed, but  
 CC the binding protein used has a higher binding constant and as it has a  
 CC low molecular weight, it is less antigenic than other proteins used in  
 CC similar procedures. The method is better at removing endotoxin than  
 CC previous methods e.g. ultrafiltration, ion exchange, affinity  
 CC chromatography etc. which cannot distinguish between proteins and  
 CC endotoxins.  
 SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11

Db 40 RRLKWKYKGF 50

#### RESULT 14

W65112 ID W65112 standard; protein; 101 AA.

AC W65112; 28-SEP-1998 (first entry)  
 DE Horseshoe crab ENP protein fragment.  
 KW Endotoxin neutralising protein; ENP; horseshoe crab; detection; biocide;  
 KW quantification; oil-in-water emulsion; coolant; metal cutting; aluminium;  
 KW steel rolling; metal forming; rust prevention; metal cleaner; antibiotic;  
 KW heat treating fluid; paint; adhesive; food production; waste water;  
 KW air quality; pharmaceutical; Gram positive bacteria; diagnostic.  
 OS Carcinoscopus rotundicauda.  
 PN WO9821357-A1.

PD 22-MAY-1998.

PF 13-NOV-1997; U20556.

PR 13-NOV-1997; US-967038.

PR 13-NOV-1996; US-030594.

PA (ASCA-) ASSOC CAPE COD INC.

PI Novitsky TJ, Ridge RJ, Sloyer JL;

DR WPI; 98-297959/26.

PT Fluorescently labelled endotoxin neutralising proteins - used for  
 PT detection of endotoxins in e.g. metal working fluids, paints,  
 PT adhesives, foods or personal care, pharmaceutical or medical  
 PT products

PS Claim 12; Page 27; 42pp; English.

CC This sequence represents an endotoxin neutralising protein (ENP)  
 CC fragment from the horseshoe crab. This sequence is used in a method  
 CC which detects or quantifies endotoxins in samples such as oil-in-water  
 CC emulsions used as coolants for metal cutting, aluminium and steel  
 CC rolling, metal forming, rust preventatives, metal cleaners, metal heat  
 CC treating fluids, paints and adhesives, food products and food handling  
 CC surfaces, waste water and cooling tower water, personal care products,  
 CC indoor and outdoor air quality, pharmaceutical products, and medical  
 CC diagnostic products, e.g. blood, urine, CSF, or lavages. In addition the  
 CC method can be used to measure molecules specific for gram positive  
 CC bacteria and fungi, as well as biocides and antibiotics.  
 SQ Sequence 101 AA;

Query Match 95.3%; Score 61; DB 1; Length 101;  
 Best Local Similarity 90.9%; Pred. No. 0.0051;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11

:|||||

Db 40 RRLKWKYKGF 50

#### RESULT 15

W65113 ID W65113 standard; Protein; 105 AA.

AC W65113; 28-SEP-1998 (first entry)

DE Horseshoe crab recombinant ENP protein fragment.

KW Endotoxin neutralising protein; ENP; horseshoe crab; detection; biocide;  
 KW quantification; oil-in-water emulsion; coolant; metal cutting; aluminium;  
 KW steel rolling; metal forming; rust prevention; metal cleaner; antibiotic;  
 KW heat treating fluid; paint; adhesive; food production; waste water;  
 KW air quality; pharmaceutical; Gram positive bacteria; diagnostic.  
 OS Carcinoscopus rotundicauda.  
 PN WO9821357-A1.

PD 22-MAY-1998.

PF 13-NOV-1997; U20556.

PR 13-NOV-1997; US-967038.

PR 13-NOV-1996; US-030594.

PA (ASCA-) ASSOC CAPE COD INC.

PI Novitsky TJ, Ridge RJ, Sloyer JL;

DR WPI; 98-297959/26.

DR N-PSDB; V35327.

PT Fluorescently labelled endotoxin neutralising proteins - used for  
 PT detection of endotoxins in e.g. metal working fluids, paints,  
 PT adhesives, foods or personal care, pharmaceutical or medical  
 PT products

PS Claim 13; Page 28; 42pp; English.

CC This sequence represents a recombinant endotoxin neutralising protein  
 CC (ENP) fragment from the horseshoe crab. This sequence is used in a method  
 CC which detects or quantifies endotoxins in samples such as oil-in-water  
 CC emulsions used as coolants for metal cutting, aluminium and steel  
 CC rolling, metal forming, rust preventatives, metal cleaners, metal heat  
 CC treating fluids, paints and adhesives, food products and food handling  
 CC surfaces, waste water and cooling tower water, personal care products,  
 CC indoor and outdoor air quality, pharmaceutical products, and medical  
 CC diagnostic products, e.g. blood, urine, CSF, or lavages. In addition the  
 CC method can be used to measure molecules specific for gram positive  
 CC bacteria and fungi, as well as biocides and antibiotics.  
 SQ Sequence 105 AA;

Query Match 95.3%; Score 61; DB 1; Length 105;  
 Best Local Similarity 90.9%; Pred. No. 0.0052;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKWKYKGF 11

:|||||

Db 44 RRLKWKYKGF 54

Search completed: September 7, 1999, 23:18:25  
 Job time: 1681 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 1999, 23:17:36 ; Search time 41.46 Seconds  
(without alignments)  
16.329 Million cell updates/sec

Title: US-09-124-280A-45

Perfect score: 64

Sequence: 1 KRLKWKYKGF 11

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL\_10.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	62.5	485	8	O78396 brachyloma
2	39	60.9	682	2	O51192 borrelia bu
3	39	60.9	2004	2	O33906 shewanella
4	39	60.9	273	3	O05782 saccharomyc
5	39	60.9	454	5	Q17511 caenorhabdi
6	39	60.9	276	13	O93287 ambystoma m
7	38	59.4	168	1	O27914 methanobact
8	38	59.4	754	5	O20418 caenorhabdi
9	37	57.8	1356	3	O42731 ascobolus i
10	37	57.8	186	3	O59678 schizosacch
11	37	57.8	1336	3	O42803 ascobolus 1
12	37	57.8	2207	3	O93845 emericeila
13	37	57.8	320	10	O92UX8 arabidopsis
14	37	57.8	1784	12	O85521 feline leuk
15	37	57.8	1786	12	O89811 feline leuk
16	36	56.2	957	1	O26426 methanobact
17	36	56.2	438	1	O58314 pyrococcus
18	36	56.2	295	1	O58667 pyrococcus
19	36	56.2	758	1	O58981 pyrococcus
20	36	56.2	445	1	O59214 pyrococcus
21	36	56.2	480	2	P96634 bacillus su
22	36	56.2	398	2	P76168 escherichia
23	36	56.2	253	2	O92DQ4 rickettsia
24	36	56.2	910	3	P87211 orpinomyces
25	36	56.2	536	4	Q13529 homo sapien
26	36	56.2	569	5	Q18211 caenorhabdi
27	36	56.2	815	5	Q18422 caenorhabdi
28	36	56.2	345	8	O78881 chanaelinor
29	36	56.2	773	12	O39274 equine herp

30	36	56.2	362	12	Q69261 equine herp
31	35.5	55.5	610	2	O32582 escherichia
32	35	54.7	558	1	O58315 pyrococcus
33	35	54.7	226	2	Q06073 bacillus me
34	35	54.7	149	2	O50557 actinobacil
35	35	54.7	595	2	O67249 aquifex aeo
36	35	54.7	329	2	O30601 bacillus su
37	35	54.7	168	2	P73576 synecocyst
38	35	54.7	337	2	P73576 synecocyst
39	35	54.7	301	2	Q92GM3 leptospira
40	35	54.7	1246	3	Q12276 saccharomyc
41	35	54.7	681	3	P87159 schizosacch
42	35	54.7	394	3	O94708 schizosacch
43	35	54.7	872	5	P90523 dictyostell
44	35	54.7	776	5	O45269 caenorhabdi
45	35	54.7	735	6	O18817 oryctolagus

#### ALIGNMENTS

RESULT 1  
O78396 PRELIMINARY; PRT; 485 AA.  
AC O78396;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE MATURASE (FRAGMENT).  
GN MATK.  
OS Brachyloma daphnoides.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asterales; Ericaceae; Brachyloma.  
[1]  
RN SEQUENCE FROM N.A.  
RA KRON K.A., FULLER R., CRAYN D.M., GADEK P.A., QUINN C.J.;  
RT "Phylogenetic relationships of epacrids and vaccinioids (Ericaceae  
s.l.) based on matk sequence data.";  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF015633; AAC62165.1; -  
DR PFAM; PF01348; Intron\_Maturas2; 1.  
KW Chloroplast.  
FT NON\_TER 485 485  
SQ SEQUENCE 485 AA; 57971 MW; 47955DDF8 CRC32;

Query Match 62.5%; Score 40; DB 8; Length 485;  
Best Local Similarity 63.6%; Pred. NO. 37;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
| | | | |  
DB 259 KDLKWKFKGPF 269

RESULT 2  
O51192 PRELIMINARY; PRT; 682 AA.  
AC O51192;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 78.5 KD PROTEIN.  
GN BB0170.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN-ATCC 35210 / B31;  
RX MEDLINE; 98065943.  
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,

RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,  
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,  
 RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
 RA UTTERBACK T., WATHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
 RA SMITH H.O., VENTER J.C.:  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 RT burgdorferi.";  
 RL Nature 390:580-586(1997).  
 DR EMBL: AE001128; AAC66568.1; --  
 DR TIGR: BB0170; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 682 AA; 78474 MW; 4233033A CRC32;

Query Match 60.9%; Score 39; DB 2; Length 682;  
 Best Local Similarity 54.5%; Pred. No. 76;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 KRLWKYKGF 11  
 :||:|:|:|:|:  
 Db 369 KKLKYEVSCEF 379

RESULT 3  
 C03906 PRELIMINARY; PRT; 2004 AA.  
 ID C03906;  
 AC C03906;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE HYPOTHETICAL 216.9 KD PROTEIN.  
 OS Shewanella sp. SCRC-2738.  
 CC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;  
 OC Shewanella.

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-SCRC-2738;  
 RC MEDLINE; 97419510.  
 RA TAKEYAMA H., TAKEIDA D., YAZAWA K., YAMADA A., MATSUNAGA T.;  
 RT "Expression of the eicosapentaenoic acid synthesis gene cluster from  
 RT Shewanella sp. in a transgenic marine cyanobacterium, Synechococcus  
 sp.";  
 RL Microbiology 143:0-0(0).  
 DR EMBL: U73935; AAB81125.1; --  
 DR PFAM: PF00109; ketoacyl-synt. 1.  
 DR PFAM: PF01377; Thioester\_dehyd. 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 2004 AA; 216919 MW; 19AAB50 CRC32;

Query Match 60.9%; Score 39; DB 2; Length 2004;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 LKWKYKGF 10  
 :|||:|:  
 Db 1944 IKWKYRGQ 1951

RESULT 4  
 Q05782 PRELIMINARY; PRT; 273 AA.  
 ID Q05782;  
 AC Q05782;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE SIMILAR TO S. CEREBISIAE HYPOTHETICAL PROTEIN IN QR15 5' REGION.  
 GN L8167.18.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-S288C (AB972);  
 RC MEDLINE; 97313267.  
 RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,  
 RA BENES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOFT A.,  
 RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEBLING U., HEUMANN K.,  
 RA HEUSS-NEITZEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,  
 RA LOUIS E.J., MESSENGUY F., MEWES H.W., MIOGA T., MOSTL D.,  
 RA MULLER-AUER S., NENTWICH U., OBERMAIER B., PIRAVANDI E., POHL T.M.,  
 RA PORTETTELLE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,  
 RA SCHARFE M., SCHERENS B., SCHOLLER P., SCHWABER C., SCHWARZ S.,  
 RA UNDERWOOD A.P., URRESTARAZU L.A., VANDENBOL M., VERHASSLT P.,  
 RA VIERENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUIT R., WEDLER E.,  
 RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOEISEL J.D.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT xii.";  
 RL Nature 387:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA PAULEY A.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA WATERSTON R.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C (AB972);  
 RA CHERRY J.M.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U14913; AAB67439.1; --  
 SQ SEQUENCE 273 AA; 32391 MW; 9EF66719 CRC32;

Query Match 60.9%; Score 39; DB 3; Length 273;  
 Best Local Similarity 60.0%; Pred. No. 31;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 RLKWKYKGF 11  
 :|||:|:  
 Db 167 RLKWKYKKNY 176

RESULT 5  
 Q17511 PRELIMINARY; PRT; 454 AA.  
 ID Q17511;  
 AC Q17511;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE B0491.4 PROTEIN.  
 GN B0491.4.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SULSTON J.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR ENBL: 249907; CAA90086.1; -.  
 SQ SEQUENCE 454 AA; 52667 MW; 25F66E2C CRC32;

Query Match 60.9%; Score 39; DB 5; Length 454;  
 Best Local Similarity 54.5%; Pred. No. 51;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11  
 :|:|:| |:  
 Db 103 ERKWKDKEG 113

RESULT 6  
 O93287 PRELIMINARY; PRT; 276 AA.  
 AC O93287;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE HOMEBOX TRANSCRIPTION FACTOR.  
 GN HOXD-11.  
 OS Ambystoma mexicanum (Axolotl).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Caudata; Salamandroidea; Ambystomatidae; Ambystoma.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95324399.  
 RA GARDINER D.M., BLUMBERG B., KOMINE Y., BRYANT S.V.;  
 RT "Regulation of HoxA expression in developing and regenerating axolotl  
 RT limbs.";  
 RL Development 121:1731-1741(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 98370927.  
 RA TOROK M.A., GARDINER D.M., SHUBIN N.H., BRYANT S.V.;  
 RT "Expression of Hoxd genes in developing and regenerating axolotl  
 RT limbs.";  
 RL Dev. Biol. 200:225-233(1998).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR ENBL: AF031481; AAC34744.1; -.  
 DR PFAM: PF00046; homeobox; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 SQ SEQUENCE 276 AA; 31673 MW; 5C829680 CRC32;

Query Match 60.9%; Score 39; DB 13; Length 276;  
 Best Local Similarity 45.5%; Pred. No. 31;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLWKYKGF 11  
 :|:|:| |:  
 Db 68 ERSKWYRGSY 78

RESULT 7  
 O27914 PRELIMINARY; PRT; 168 AA.  
 AC O27914;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
 DE CONSERVED PROTEIN.  
 GN MTH1892.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacterium.

RN SEQUENCE FROM N.A.  
 RP STRAIN-DELTA H;  
 RX MEDLINE; 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTIER B., QIU D.,  
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR ENBL: AE00940; AAB86352.1; -.  
 SQ SEQUENCE 168 AA; 18891 MW; DFB942B6 CRC32;

Query Match 59.4%; Score 38; DB 1; Length 168;  
 Best Local Similarity 55.6%; Pred. No. 28;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LKWKYKGF 11  
 |:::|:  
 Db 96 LSWKRFGEF 104

RESULT 8  
 Q20418 PRELIMINARY; PRT; 754 AA.  
 AC Q20418;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-JAN-1999 (Tremblrel. 09, Last annotation update)  
 DE F44G4.1 PROTEIN.  
 GN F44G4.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SIMS M.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR ENBL: Z49910; CAA90124.1; -.  
 DR ENBL: Z70034; CAA90124.1; JOINED.  
 SQ SEQUENCE 754 AA; 87875 MW; 3739D9F7 CRC32;

Query Match 59.4%; Score 38; DB 5; Length 754;  
 Best Local Similarity 70.0%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLKWKYKGF 11  
 |:::|:  
 Db 720 RLKWLQKGF 729

RESULT 9  
 O42731 PRELIMINARY; PRT; 1356 AA.  
 AC O42731;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE CS-DNA-METHYLTRANSFERASE.  
 GN MASC2.  
 OS Ascombolus immersus.  
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pezizales; Ascombolaceae;  
 OC Ascombolus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-RN42;  
 RX MEDLINE; 98121243.

RA CHERNOV A.V., VOLLMEYER P., WALTER J., TRAUTNER T.A.;  
 RT "Masc2, a CS-DNA-methyltransferase from Ascobolus immersus with  
 RL similarity to methyltransferases of higher organisms.";  
 RL Bio1. Chem. 378:1467-1473(1997).  
 DR EMBL; AF030976; AAC03766.1; -;  
 DR PFAM; PF01426; BAH; 1;  
 DR PFAM; PF00145; DNA\_methylase; 3;  
 KW Transferase; Methyltransferase;  
 SQ SEQUENCE 1356 AA; 153536 MW; AAA066BD CRC32;

Query Match 57.8%; Score 37; DB 3; Length 1356;  
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKG 9  
 |||||:|  
 Db 387 KRLKWEIRG 395

RESULT 10  
 ID O59678 PRELIMINARY; PRT; 166 AA.  
 AC O59678;  
 DT 01-JAN-1999 (TRENBLrel. 09, Created)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last sequence update)  
 DT 01-JAN-1999 (TRENBLrel. 09, Last annotation update)  
 DE HYPOTHETICAL 19.0 KD PROTEIN C29A3.16 IN CHROMOSOME II.  
 GN SPBC29A3.16.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO YEAST YOR294W AND HUMAN KIA0112.  
 DR EMBL; AL022299; CAA18393.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 166 AA; 19044 MW; 7445291A CRC32;

Query Match 57.8%; Score 37; DB 3; Length 166;  
 Best Local Similarity 85.7%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 KWKYKKG 10  
 |||||  
 Db 123 KWKYKKG 129

RESULT 11  
 ID O42803 PRELIMINARY; PRT; 1336 AA.  
 AC O42803;  
 DT 01-JUN-1998 (TRENBLrel. 06, Created)  
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE DNA-CS-METHYLTRANSFERASE (EC 2.1.1.37) (DNA  
 DE (CYTOSINE-5-)-METHYLTRANSFERASE).  
 GN MASC2.  
 OS Ascobolus immersus.  
 OC Plasmid pCG92.  
 OC Ascobolus.  
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Pezizales; Ascombolaceae;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RN42;  
 RX MEDLINE; 98182583.  
 RA GOYON C.;  
 RT "Sequence and gene content in 35 kb genomic clone mapping in the

RT human Xq27.1 region.";  
 RL DNA Seq. 8:1-4(1997).  
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA =  
 CC S-ADENOSYL-L-HOMOCYSTEINE + DNA CONTAINING 5-METHYLCYTOSINE.  
 DR EMBL; Z96933; CAB09661.1; -;  
 DR PFAM; PF01426; BAH; 1;  
 DR PFAM; PF00145; DNA\_methylase; 3;  
 KW Transferase; Methyltransferase; Plasmid.  
 SQ SEQUENCE 1336 AA; 151125 MW; A2042DCF CRC32;

Query Match 57.8%; Score 37; DB 3; Length 1336;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRLKWKYKG 9  
 |||||:|  
 Db 387 KRLKWEIRG 395

RESULT 12  
 ID O93845 PRELIMINARY; PRT; 2207 AA.  
 AC O93845;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE EPSILON HOMOLOG.  
 GN NIMP.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Plecomycetes;  
 OC Eurotiales; Trichomycetes; Emericella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA JAMES S.W., CRAWFORD G.E., WEXLER A.N.;  
 RL "himp DNA Pol Epsilon gene of Aspergillus nidulans";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF019254; AAD01637.1; -;  
 SQ SEQUENCE 2207 AA; 252648 MW; 14A06363 CRC32;

Query Match 57.8%; Score 37; DB 3; Length 2207;  
 Best Local Similarity 45.5%; Pred. No. 5.1e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KRLKWKYKGF 11  
 :|||:|:|:|  
 Db 648 RRLPWAIRGEF 658

RESULT 13  
 ID O92UX8 PRELIMINARY; PRT; 320 AA.  
 AC O92UX8;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 DE F15K20.13 PROTEIN.  
 GN F15K20.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RC ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,  
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;  
 RL "Arabidopsis thaliana chromosome II BAC F15K20 genomic sequence";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005824; AAC73023.1; -;  
 SQ SEQUENCE 320 AA; 36096 MW; C0DC45B6 CRC32;

Query Match 57.8%; Score 37; DB 10; Length 320;  
Best Local Similarity 55.6%; Pred. No. 77;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KRLWKYK 9  
|||:|:|  
Db 240 KRLQWNG 248

## RESULT 14

O85521 PRELIMINARY; PRT; 1784 AA.  
AC Q85521;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE GAG-POL PRECURSOR POLYPROTEIN GPR80.  
OS Feline leukemia virus.  
OC Viruses; Retrovirdae;  
OC Mammalian type C retroviruses.  
[1]  
"RP SEQUENCE FROM N.A.  
"RC STRAIN-FELV-FAIDS;  
RX MEDLINE; 88119207.  
RA DONAHUE P.R., HOOVER E.A., BELTZ G.A., RIEDEL N., HIRSCH V.M.,  
RA OVERBAUGH J., MULLINS J.I.;  
RT "Strong sequence conservation among horizontally transmissible,  
RT minimally pathogenic feline leukemia viruses.";  
RL J. Virol. 62:722-731(1998).  
DR EMBL; M18247; AAA93092.1; -  
DR PFAM; PF01140; gag\_MA; 1.  
DR PFAM; PF01141; gag\_P12; 1.  
DR PFAM; PF00552; integrase; 1.  
DR PFAM; PF00075; rnsase; 1.  
DR PFAM; PF00065; rve; 1.  
DR PFAM; PF00077; rvp; 1.  
DR PFAM; PF00078; rvt; 1.  
DR PFAM; PF00079; rvt; 1.  
DR PFAM; PF00080; rvt; 1.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
KW Signal; Polyprotein; Hydrolase; Aspartyl protease.  
FT SIGNAL 1766 1784  
FT CHAIN 75 576  
FT CHAIN 75 576  
FT CHAIN 202 271  
FT CHAIN 272 519  
FT CHAIN 520 576  
FT CHAIN 577 1784  
FT CHAIN 577 701  
FT CHAIN 702 1368  
FT CHAIN 1369 1784  
SQ SEQUENCE 1784 AA; 200157 MW; 003D5695 CRC32;

Query Match 57.8%; Score 37; DB 12; Length 1784;  
Best Local Similarity 60.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KRLWKYK 10  
|||:|:|  
Db 1401 KRGWEYRGK 1410

## RESULT 15

O89811 PRELIMINARY; PRT; 1786 AA.  
AC O89811;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
DE GAG-POL PRECURSOR POLYPROTEIN GPR80.  
OS Feline leukemia virus.

OC Viruses; Retrovirdae; Retroviridae;  
OC Mammalian type C retroviruses.  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RICKARD SUBGROUP A;  
RX MEDLINE; 98362106.  
RA CHEN H., BECHTEL M.K., SHI Y., PHIPPS A., MATHES L.E., HAYES K.A.,  
RA ROY-BURMAN P.;  
RT "Pathogenicity induced by feline leukemia virus, Rickard strain,  
RT subgroup A plasmid DNA (pFRA).";  
RL J. Virol. 72:7048-7056(1998).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RICKARD SUBGROUP A;  
RA CHEN H., ROY-BURMAN P.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF052723; AAC31801.1; -  
DR PFAM; PF01140; gag\_MA; 1.  
DR PFAM; PF01141; gag\_P12; 1.  
DR PFAM; PF00552; integrase; 1.  
DR PFAM; PF00075; rnsase; 1.  
DR PFAM; PF00065; rve; 1.  
DR PFAM; PF00077; rvp; 1.  
DR PFAM; PF00078; rvt; 1.  
DR PFAM; PF00079; rvt; 1.  
DR PFAM; PF00080; rvt; 1.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
KW Polyprotein; Hydrolase; Aspartyl protease.  
FT CHAIN 75 576  
FT CHAIN 75 576  
FT CHAIN 202 271  
FT CHAIN 272 519  
FT CHAIN 520 576  
FT CHAIN 577 1786  
FT CHAIN 577 701  
FT CHAIN 702 1370  
FT CHAIN 1371 1786  
SQ SEQUENCE 1786 AA; 200201 MW; 93530DF1 CRC32;

Query Match 57.8%; Score 37; DB 12; Length 1786;  
Best Local Similarity 60.0%; Pred. No. 4.2e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KRLWKYK 10  
|||:|:|  
Db 1403 KRGWEYRGK 1412

Search completed: September 7, 1999, 23:17:38  
Job time: 1781 sec

**THIS PAGE BLANK (USPTO)**